Essays on Courts, Randomization, and Experiments

Dane Ross Thorley

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
in the Graduate School of Arts and Sciences

COLUMBIA UNIVERISTY

2019

Abstract

Essays on Courts, Randomization, and Experiments

Dane Ross Thorley

This dissertation comprises three chapters that explore and expand on the use of experimentation and randomization in the study of courts, judges, and the law:

Chapter 1: This Chapter reviews the two most prominent procedural approaches to addressing judicial conflicts of interest in U.S. courts—judicial self-recusal and in-court disclosure. These procedural approaches fail to account for the legal and institutional dynamics that surround the relationship between judges, attorneys, and the adjudicative process. I argue that judges do not recuse themselves, that attorneys will not ask them to, and that if we understand both the legal and extra-legal incentives at play in these decisions, this should not surprise us. The shortcomings of recusal and disclosure are particularly salient in the context of judicial campaign finance, where judges often face the acute dilemma of being assigned to preside over cases in which one of the parties or attorneys has contributed to their election campaign.

To support these claims, Chapter 1 presents the results of a randomized field experiment which I identify active Wisconsin and Texas civil cases that feature donor-attorneys. The experiment randomly assigns a portion of the judges presiding over these cases to receive a letter from an NGO identifying the potential conflict and requesting recusal. The empirical results support the growing skepticism surrounding judicial self-recusal and raise doubts that judicial disclosure is an efficacious remedy. Building on these results, the Chapter explores two potential alternatives—one procedural and one institutional—that better account for the realities of judicial conflicts of interest and the incentives of court actors.

Chapter 2: This Chapter contributes to the growing literature challenging the general assumption of and reliance on random judicial assignment by identifying common court procedures and practices that threaten unbiased causal inference. These "de-randomizing" events, including differing probabilities of assignment, post-assignment judicial changes, non-random missingness, and non-random assignment itself, should be accounted for when making causal claims but are commonly either ignored or not even recognized by researchers utilizing random judicial assignment. The Chapter explores how these de-randomizing events violate the key empirical assumptions underlying randomized studies and offers methodological solutions and presents original data from a survey of the 30 largest U.S. state-level criminal courts, outlining their assignment protocols and identifying the extent to which they feature the de-randomizing events described.

Chapter 3: In *Williams-Yulee v. The Florida Bar* (2015), the Supreme Court ruled that a Florida law banning direct campaign solicitation by judicial candidates was not a violation of the First Amendment. In doing so, the majority relied on several untested empirical claims, including the assertion that direct solicitation has a distinctly stronger impact on the public's confidence in the judiciary than indirect solicitation. This chapter provides a short but focused evaluation of these empirical claims. A nationally-representative survey experiment presents subjects with a hypothetical vignette in which a state trial-level judge runs for election and utilizes one of various campaign fundraising tactics. The survey then presents subjects with questions relating to the trust and legitimacy that they associate with both the judicial system presented in the vignette and their actual state- and federal-level government institutions. The results suggest that the public does not discern any significant difference between direct and indirect judicial solicitation but does see other judicial campaign features (promises of recusal and the amount of the donations) as salient in regard to trust and legitimacy. These findings are at odds with the empirical assumptions that the majority relied upon in the *Williams-Yulee*

decision and highlight the value that survey experiments can play in evaluating empirical claims made by the Supreme Court.

Contents

List of Tables

Chapter 1

Chapter 2

Chapter 3

Appendix A (Chapter 1)

Appendix B (Chapter 3)

List of Figures

Chapter 1

Chapter 2

Appendix A (Chapter 1)

Acknowledgments

This dissertation is dedicated to the three most important people in my life. To Brit, who stayed in Provo while I finished my undergrad, and then moved to New York with me for more school, and then moved to New Haven for more school, and then moved to Las Vegas for a clerkship, and will soon move back to Provo for my new job. I'm indebted to you and look forward to shifting the focus back to your professional aspirations over the next few years. Marrying you was the best decision of my life, and I'll be forever thankful that you stuck with me through all this. And to my two children, Jack and Sloane, you certainly didn't make this dissertation any easier to write, but your unconditional love over these last few years—despite my absence (physical and mental) at times—has been the brightest of lights.

Of course, I owe a lot to my parents, Kennedy and Steve, who found the perfect balance of love, expectation, and support for me. It's impossible to describe in words how much I love and respect you. I'm also grateful to my other family members, including my in-laws, for all their support.

There are many who have mentored me during my academic journey. To Dan Nielson and Mike Findley, thank you for hiring me ten years ago, introducing me to political science and field experiments, and laying the groundwork for my intellectual identity. To Adam Chilton, Ben Liebman, Bob Scott, Ian Ayres, Josh Mitts, Katarina Pistor, Rick Brooks, and Tom Tyler, thank you for helping me become a legal academic. To Jeff Lax, thank you for introducing me to judicial politics and making sure I didn't become too much of a legal academic. And to Don Green, thank you for taking me under your wing, thank you for your patient guidance, and thank you for making Brit and I part of your family when we moved to New York.

I am also grateful to all those who provided non-academic support in those moments where I was able to step away from school. To the Morningside Heights and New Haven Congregations, thank you for fostering an environment where I could critically (and sometimes heretically) engage with my faith and relationship with God. To my PhD cohort and law classmates, thank you for your companionship. To the Yale Grad Rugby Team, thank you for getting me out of my office and reminding me that I'm not 20 anymore. To Judge Gordon and Kara Rickey, thank you for putting up with a clerk who was on the job market and working on his dissertation. And to all the other friends that I've made over the last seven years, thank you for making me a part of your life.

Finally, I want to recognize all the many friends, teachers, colleagues, conference attendees, and classmates who have provided feedback on the three studies included in this dissertation and any of the other work I somehow put together over the last seven years.

Introduction

Experimental research design is widely regarded as the gold standard for making causal claims through empirical research. Recognizing this, social scientists are increasingly utilizing randomization and experimentation to explore an ever-expanding array of questions and theories.[1] Political scientists and legal scholars who study courts, however, have yet to fully adopt this methodological innovation.[2] Field experiments involving judges or conducted in a courtroom context are nearly non-existent. Naturally occurring randomization in the court—predominantly the random assignment of judges—is exploited in a growing number of studies but is often used incorrectly. And while survey experiments have occasionally been used to study judges, they are rarely used to investigate and critique the assumptions underlying the legal doctrines that U.S. courts generate. This dissertation advances the use of experiments and randomization in the courtroom through three essays addressing three distinct methodological techniques.

Chapter 1 features the first ever randomized field experiment conducted on judges presiding over active cases. It begins by reviewing the two most prominent procedural approaches to addressing judicial conflicts of interest in U.S. courts—judicial self-recusal and in-court disclosure. These procedural approaches fail to account for the legal and institutional dynamics that surround the relationship between judges, attorneys, and the adjudicative

[1] *See* James N. Druckman, Donald P. Green, James H. Kuklinski, & Arthur Lupia, *Experimentation in Political Science* in CAMBRIDGE HANDBOOK OF EXPERIMENTAL POLITICAL SCIENCE 3-14 (James N. Druckman, Donald P. Green, James H. Kuklinski, and Arthur Lupia, eds., 2010); Steven D. Levitt & John A. List, *Field Experiments in Economics: The Past, Present, and Future*, 53 EURO. ECON. REV. 1 (2009).

[2] *See* Donald P. Green & Dane R. Thorley, *Field Experimentation and the Study of Law and Policy*, 10 ANN. REV. L. & SOC. SCI. 53 (2014); Dane Thorley and Jacob Kopas, *Experiments in the Court: The Legal and Ethical Challenges of Running Randomized Field Experiments in the Courtroom*, Working Paper (2018), available at https://ssrn.com/abstract=2994298.

process. I argue that judges do not recuse themselves, that attorneys will not ask them to, and that if we understand both the legal and extra-legal incentives at play in these decisions, this should not surprise us. The shortcomings of recusal and disclosure are particularly salient in the context of judicial campaign finance, where judges often face the acute dilemma of being assigned to preside over cases in which one of the parties or attorneys has contributed to their election campaign.

To support these claims, Chapter 1 presents the results of a randomized field experiment which I identify active Wisconsin and Texas civil cases that feature donor-attorneys. The experiment randomly assigns a portion of the judges presiding over these cases to receive a letter from an NGO identifying the potential conflict and requesting recusal. The empirical results support the growing skepticism surrounding judicial self-recusal and raise doubts that judicial disclosure is an efficacious remedy. Building on these results, the Chapter explores two potential alternatives—one procedural and one institutional—that better account for the realities of judicial conflicts of interest and the incentives of court actors.

Chapter 2 is a methodological essay that contributes to the growing literature challenging the general assumption of and reliance on random judicial assignment by identifying common court procedures and practices that threaten unbiased causal inference. These "de-randomizing" events, including differing probabilities of assignment, post-assignment judicial changes, non-random missingness, and non-random assignment itself, should be accounted for when making causal claims but are commonly either ignored or not even recognized by researchers utilizing random judicial assignment. The Chapter explores how these de-randomizing events violate the key empirical assumptions underlying randomized studies and offers methodological solutions and presents original data from a survey of the 30 largest U.S. state-level criminal courts, outlining their assignment protocols and identifying the extent to which they feature the de-randomizing events described.

Chapter 3 features a randomized survey experiment that assesses the key empirical assumptions relied upon by U.S. Supreme Court Justices in *Williams-Yulee v. The Florida Bar* (2015). In that case the Supreme Court ruled that a Florida law banning direct campaign solicitation by judicial candidates was not a violation of the First Amendment. In doing so, the majority relied on several untested empirical claims, including the assertion that direct solicitation has a distinctly stronger impact on the public's confidence in the judiciary than indirect solicitation. This chapter provides a short but focused evaluation of these empirical claims. A nationally-representative survey experiment presents subjects with a hypothetical vignette in which a state trial-level judge runs for election and utilizes one of various campaign fundraising tactics. The survey then presents subjects with questions relating to the trust and legitimacy that they associate with both the judicial system presented in the vignette and their actual state- and federal-level government institutions. The results suggest that the public does not discern any significant difference between direct and indirect judicial solicitation but does see other judicial campaign features (promises of recusal and the amount of the donations) as salient in regard to trust and legitimacy. These findings are at odds with the empirical assumptions that the majority relied upon in the *Williams-Yulee* decision and highlight the value that survey experiments can play in evaluating empirical claims made by the Supreme Court.

Chapter 1[*]

Why Judges Don't Recuse Themselves and Attorneys Don't Ask Them To: A Field Experiment Testing the Efficacy of Recusal and Disclosure

[*] The experimental study in this Chapter has received IRB approval and was funded by grants from the Open Society Foundation, the Democracy Fund, and the Columbia University School of Arts and Sciences, none of which is responsible for the content of this Chapter.

Portions of the experimental analysis were conducted as a collaborative effort with Jonathan S. Krasno (Binghamton University), Donald P. Green (Columbia University), Costas Panagopoulos (Northeastern University), and Michael Schwam-Baird (Columbia University). The results of these portions of the analysis will also be presented in a preceding, coauthored article, *Please Recuse Yourself: A Field Experiment Exploring the Relationship Between Campaign Donations and Judicial Recusal.*

Note: Although this chapter is a solo-authored piece, in order to reflect the contributions of my colleagues in designing and implementing portions of the field experiment, this Chapter will use the pronoun "we."

A pre-analysis plan for the experimental design and analysis featured in this Chapter is registered with Evidence in Governance and Politics and can be found at http://egap.org/registration/719. Any deviations from that pre-analysis plan are noted in the accompanying text.

"If you point the barrel of a recusal motion at a Texas judge, make sure it is loaded with a silver bullet." – Common saying among attorneys in Texas[3]

I. Introduction

In theory, judicial recusal is a simple and effective procedural remedy for judicial conflicts of interest—the judge will be on the lookout for factors in a case that might prevent her from ruling impartially, and if she identifies any, she can remove herself and have the case assigned to a more neutral adjudicator. As a safeguard, the judge is expected to disclose potential conflicts so that the parties and attorneys who might be disadvantaged can then push for recusal when necessary. These behavioral assumptions undergird the current legal regime of judicial recusal and disqualification in almost all U.S. courts. However, these assumptions are likely false. Working within a legal framework that allows for but does not explicitly require recusal under most circumstances, judges are unlikely to remove themselves because judges do not like to recuse: doing so requires a recognition of the conflict, takes time, may result in lost knowledge of a case, and is often perceived as a dereliction of judicial duties.

With an increasing understanding of the role that these extra-legal factors play in the recusal decisions of judges, those skeptical of existing recusal laws have suggested that an increased focus on in-court disclosure of conflicts can serve as a simple and efficacious procedural complement. By informing all the participants in a case of a potential conflict of interest, the recusal decision can be transferred, in part, to the individuals who are more incentivized to identify partiality and seek removal—namely the parties and attorneys against whom the conflict lies. Attorneys, however, are faced with their own set of extra-legal incentives that discourage them from asking for recusal when necessary: they know that the decision to

[3] Jeff Nobles, *Judicial Recusal and Attorney Disqualification: An Ethic for Litigators & Other Aliens in a Strange Land*, Texas Bar College of Legal Education: Advanced Civil Appellate Courses (1999), available at http://www.texasbarcle.com/Materials/Events/1295/66861_01.pdf.

recuse is almost entirely up to the personal discretion of the judge, and they do not want to risk the repercussions of impugning the judge's character in the likelihood that the judge will not recuse and the near certainty that they will be back in front of the judge again for future cases.

These behavioral considerations are particularly salient in the context of judicial campaign finance, where judges often face the acute dilemma of being assigned to preside over cases in which one of the parties or attorneys has contributed to the judge's election campaign. Empirical work is increasingly evidencing a strong causal relationship between whether a court participant had donated to a judge and how well that participant fares in the case, and a majority of both the public and the legal community (including the judges themselves) believe that campaign donations are unfairly influencing judicial decisions. If recusal and disclosure procedures are efficacious in addressing conflicts of interest, judges should be recusing in some, if not many, of the cases that feature campaign donors. Yet the nature of judicial elections and campaign finance is such that both the cost to judges of recusal and the cost to attorneys of asking the judges to recuse is likely even higher than usual.

This Chapter evaluates modern recusal procedures and practices, with a particular focus on how well they address potential conflicts in cases where an elected judge has received a campaign contribution from either a party to the case or an attorney involved in the case. It makes the following four contributions:

First, it outlines and consolidates the incentive structures behind both the judge's decision to recuse and the attorney's decision to request recusal. Although the last decade has produced a litany of scholarship exploring the judge's decision to recuse, the literature has spent relatively little time on the decision of the attorney to seek recusal, particularly in the context of state trial-level courts. Understanding the behavioral incentives of all the participants in the courtroom is essential to developing legal policy that adequately addresses its intended purpose, and this Chapter ultimately concludes that judicial self-recusal, in almost all its forms, is

unlikely to mitigate real or perceived judicial bias, even when supplemented by disclosure and transparency.

Second, this Chapter provides much-needed empirical evidence supporting the growing skepticism surrounding the current judicial recusal regime in U.S. courts. Although the *Caperton v. Massey*[4] ruling produced a veritable explosion in legal scholarship exploring (and usually deriding) the current state of recusal law in relation to judicial campaign finance, almost none of this analysis provides any empirical support for its conclusions. The data in this Chapter come from a randomized field experiment (also known as a randomized controlled trial, or RCT)—the first ever conducted on judges presiding over active cases. In the experiment, we identify active civil cases in the Wisconsin and Harris County, Texas trial courts in which one of the attorneys had donated to the judge's previous political campaign. We then randomly assign half of these judges to receive a letter identifying the potential conflict and asking the judge to recuse. The results show that judges almost never recuse themselves from these sorts of cases, even when they receive outside pressure to do so.

Third, the design of the experiment also allows us to explore the efficacy of judicial disclosure to non-donor parties as a supplement to current recusal procedures. Increased judicial disclosure of campaign contributions is among the most common remedies suggested in the literature on judicial conflicts of interest, both from those who support the current regime and those who criticize it. Those who are comfortable with the status quo assume that the current disclosure rules ensure that parties in a case are informed regarding potential sources of judicial bias. Those who criticize the current regime believe that if disclosure rates could be increased, the potentially disadvantaged attorneys would request recusal when appropriate. Although both perspectives are informed by reasonable behavioral assumptions, neither has much empirical support. The experimental results show that without our experimental

[4] Caperton v. A.T. Massey Coal Co., 129 S. Ct. 2252 (2009).

intervention, on-the-record disclosure of campaign donations is virtually nonexistent—in the over 300 cases we track during the study, we did not identify a single disclosure. However, 34 percent of the judges who received the letter requesting recusal disclosed the contribution on the court record, allowing us to then measure the downstream impact of disclosure on both requests for recusal and recusal itself. Contrary to the predictions of the literature, increased disclosure rates had no discernable effect on either the propensity for attorneys to move for recusal or for the judge to eventually recuse herself.

Lastly, building on the empirical results of the experiment (but being careful to account for its limitations), this Chapter explores potential solutions to the problems caused by donors in the courtroom. It looks first at common procedural solutions, arguing that neither automatic *per se* recusal nor independent recusal review are likely to be tenable. This Chapter then posits that the problem of money in the courtroom is best ameliorated by the combination of limited no-cause peremptory challenges paired with mandatory disclosure by the court system, as opposed to disclosure made by judge herself. It concludes by briefly exploring broad institutional reform and argues that first-order solutions such as the elimination of judicial elections or bans on judicial campaign fundraising are non-starters in the current political environment, although anonymized donations in judicial campaigns may be both feasible and effective.

This Chapter proceeds as follows. Part II reviews recusal rules and procedures, discusses the threat that judicial campaign fundraising may have on judicial impartiality and legitimacy, and then examines recusal specifically in the context of campaign finance. Part III details the incentive structure behind the judge's decision to recuse, explores disclosure as a potential solution to non-recusal, and details the incentive structure behind an attorney's decision to move for recusal. Part IV describes the randomized experiment, and the resulting data are

presented and analyzed in Part V. Part VI offers potential alternatives to the current system. Part VII concludes.

II. Judicial Conflicts of Interest, Recusal, and Campaign Finance

In this Part, we provide background on the recusal rules and procedures most commonly used in U.S. courts, with a particular focus on how they address judicial conflicts of interest stemming from campaign contributions. Subpart A briefly reviews recusal more broadly. Subpart B explores the potential effects that judicial campaign contributions have on judicial behavior and public perceptions of the judiciary, including a review of the empirical studies that have been conducted in this area. Subpart C then outlines how states with judicial elections have engineered their recusal regimes to address those potential effects.

A. Recusal in U.S. Courts

Although the U.S. legal system has devised methods for a judge to avoid conflicts of interest before they reach the bench (screening procedures, preemptive elimination of the source of bias, etc.),[5] it is inevitable that potentially problematic cases will be assigned to judges, so there must be a systematic procedural and ethical approach for dealing with such cases. For nearly every judicial body in the United States, judicial recusal is the primary solution and "perhaps the States' most reliable weapon for maintaining both the reality and the appearance of a fair hearing in a fair tribunal for every litigant."[6]

Simply defined, judicial recusal is the process through which a judge relinquishes or is disqualified from a case because of a perceived or actual bias.[7] Under most circumstances,

---

[5] One prominent approach to preemptive elimination of bias is to prevent judicial candidates from making direct campaign solicitations. The constitutionality and effectiveness of these restrictions was at issue in *Williams Yulee v. Florida Bar,* the most recent judicial election case before the U.S. Supreme Court (Williams-Yulee v. Fla. Bar, 135 S. Ct. 1656 (2015)).

[6] Brief of the Conference of Chief Justices as Amicus Curiae in Support of Neither Party, Caperton v. A.T. Massey Coal Co., No. 08-22 (U.S. Jan. 5, 2009), at 16 (internal quotations omitted).

[7] In many jurisdictions, there is a technical distinction between recusal and disqualification, with the former being a decision that is up to the judge and the latter being a decision that is imposed upon the

judges are responsible for recusing themselves *sua sponte*, although in many cases the recusal evaluation is prompted by a motion for recusal filed by the litigants or an interested third party. In this Chapter, we distinguish between for-cause recusal and peremptory challenges, the latter of which are discussed in *infra* Part VI as a potential remedy for the shortcomings of recusal.

In their current manifestations, federal and state judicial recusal requirements in the United States stem from three distinct legal sources: constitutional rights of due process, statutory and procedural rules, and judicial ethics codes. At the federal level, there is no explicit constitutional basis for judicial recusal, but the Supreme Court has found that the Due Process Clause requires recusal when "the probability of actual bias on the part of the judge or decisionmaker is too high to be constitutionally tolerable."[8] Although this somewhat circular test does not provide a clear threshold at which the probability of bias becomes unconstitutional, the Court has provided examples of violations in some of its decisions, most prominently in *Caperton v. Massey*.[9] A few state constitutions do have provisions that call for recusal under certain circumstances,[10] but the vast majority of state constitutions invoke recusal only through

judge (see, e.g. Goolsby v. State, 914 So. 2d 494, at n.1 (Fla. App. 2005); State v. Desmond, 2011 Del. Super. LEXIS 6, at 23-24 n. 62 (2011)). In practice, however, this distinction is not particularly clear (*see* ABA Model Code of Judicial Conduct, Rule 2.11, Comment [1] (2011) ("In many jurisdictions, the term "recusal" is used interchangeably with the term "disqualification.""); Note, *Caesar's Wife Revisited - Judicial Disqualification After the 1974 Amendments*, 34 WASH. & LEE L. REV. 1201 n.5 (1977)), and so for the remainder of this chapter we will refer to both processes as recusal unless the distinction plays a specific role in our argument. This is a common approach in the recusal literature (*see* RICHARD E. FLAMM, JUDICIAL DISQUALIFICATION: RECUSAL AND DISQUALIFICATION OF JUDGES 4 (3rd ed. 2017)).

[8] Caperton v. Massey, at 872, quoting Withrow v. Larkin, 421 U.S. 35, 95 S. Ct. 1456 (U.S. Apr. 16, 1975).

[9] Specifically in regard to campaign donations, the court found that the requisite probability of bias exists when "a person with a personal stake in a particular case had a significant and disproportionate influence in placing the judge on the case by raising funds or directing the judge's election campaign when the case was pending or imminent." (Caperton v. Massey, at 884).

[10] See RICHARD E. FLAMM, JUDICIAL DISQUALIFICATION: RECUSAL AND DISQUALIFICATION OF JUDGES, (3rd ed. 2017), at 31-104 (providing a comprehensive overview of the recusal approaches of all 50 states). Maryland's constitution states that, "[n]o Judge shall sit in any case wherein he may be interested, or where either of the parties may be connected with him, by affinity or consanguinity, within such degrees as now are, or may hereafter be prescribed by Law, or where he shall have been of counsel in the case." (Md. Const., Art. IV, section 7). Mississippi's constitution states that, "No judge of any court shall preside on the trial of any cause, where the parties or either of them, shall be connected with him by affinity or consanguinity, or where he may be interested in the same, except by the consent of the judge and of the parties." (Ms. Const. Article 6, section 165). Similar language exists in the New Mexico Constitution (N.M.

their own due process requirements, interpretations of which have varied depending on the state but are generally similar to the form and function of the federal Due Process Clause.[11]

The majority of federal recusal disputes are controlled by 28 U.S.C. § 455.[12] 455(b) specifies that judges must recuse when certain conditions exist, including: personal knowledge regarding the facts of the case, previous employment dealing with the case, financial interest in the case, or a familial relationship (to the third degree) with a participant in the case. 455(a) outlines the broad provision of the rule, stipulating that "[a]ny justice, judge or magistrate judge of the United States shall disqualify himself in any proceeding in which his impartiality might reasonably be questioned."[13] The appearance-of-bias standard in this "catchall" provision reflects the most common approach to regulating non-itemized sources of bias in U.S. recusal regimes.

Like the federal courts, most states have statutory regulation on when and why a judge should recuse from a case, although there are a handful of states that rely solely on constitutional provisions or court rules to address removal.[14] These statutes generally provide a list of *per se* recusal rules, although the circumstances that require recusal vary from state to state. States also vary on the standard used to make determinations regarding potential sources of bias that fall outside of the explicitly listed categories, with twenty-four states using

---

Const. Art VI, section 18), Tennessee Constitution (Tenn. Const. Art. 6, section 11), and Texas Constitution (Tex. Const. Article V, section 11).

[11] *See* Melinda A. Marbes, *Reforming Recusal Rules: Reassessing the Presumption of Judicial Impartiality in Light of the Realities of Judging and Changing the Substance of Disqualification Standards to Eliminate Cognitive Errors*, 7 ST. MARY'S J. ON LEGAL MALPRACTICE & ETHICS 238 (2016-2017), at 287.

[12] Marbes explains that the other two federal statutory venues for recusal, 28 U.S.C. sections 47 and 144, are applicable only to certain cases or have the difficult-to-attain standard of actual bias and are therefore not widely used (Melinda A. Marbes, *Reforming Recusal Rules: Reassessing the Presumption of Judicial Impartiality in Light of the Realities of Judging and Changing the Substance of Disqualification Standards to Eliminate Cognitive Errors*, 7 ST. MARY'S J. ON LEGAL MALPRACTICE & ETHICS 238 (2016-2017), at ft. 207).

[13] 28 U.S.C. § 455 (2002).

[14] *See* RICHARD E. FLAMM, JUDICIAL DISQUALIFICATION: RECUSAL AND DISQUALIFICATION OF JUDGES, (3rd ed. 2017), at 31-104.

appearance of partiality, five states requiring a showing of actual bias on the part of the judge, and the remaining states adopting a standard that incorporates both the appearance and actual bias approaches or falls somewhere in between.[15]

Additionally, both federal and state judges are ostensibly governed by additional recusal rules outlined in the respective court rules and ethics codes, although it is generally unclear to what extent these sources have the force of law and subsequently allow for justiciable claims when not followed. Although it has no force of law, the ABA Model Code of Judicial Conduct has long been the leading influence on these provisions; the Code of Conduct for U.S. judges is practically identical to the ABA Model Code, and forty-nine states have used the 1972 edition in creating their own codes.[16]

B. The (Potential) Impacts of Judicial Campaign Finance

Of the sources of conflicts facing judges, judicial campaign finance is likely the most astonishing. In almost any other context, the transfer of money from any potential litigant or attorney to an active judge is seen as an explicit threat to judicial impartiality.[17] But in the majority of judicial elections, citizens, businesses, and even active lawyers are transferring thousands, and sometimes millions, of dollars to judges through direct contributions and independent expenditures.[18]

---

[15] *See* Melinda A. Marbes, *Reforming Recusal Rules: Reassessing the Presumption of Judicial Impartiality in Light of the Realities of Judging and Changing the Substance of Disqualification Standards to Eliminate Cognitive Errors*, 7 ST. MARY'S J. ON LEGAL MALPRACTICE & ETHICS 238, 287 (2016-2017).

[16] Debra Lyn Bassett, *Judicial Disqualification in the Federal Appellate Courts,* 87 IOWA L. REV. 1213, 1229-30 (2002).

[17] To quote popular political comedian John Oliver, "Think about it. Giving money to judges wouldn't be acceptable in a state fair squash growing competition." *See also* the American Bar Association's Model Code of Judicial Conduct Canon 3.13 Comment (which states that, "Whenever a judge accepts a gift or other thing of value without paying fair market value, there is a risk that the benefit might be viewed as intended to influence the judge's decision in a case.") However, the ABA Model Code does allow the receipt of certain items, including valuable items, but requires that the judge report such occurrences (*Id.* at Canon 3.13 and 3.15).

[18] As it currently stands, judicial elections occur in the vast majority of states, although the form of election and the level of courts for which it is employed varies. 33 of 50 states employ judicial elections to initially select their judges at some level, with 33 using elections to select trial court judges, 19 using them

And yet, in elections, where judges are required to garner the popular vote in order to attain or stay in office, campaign fundraising is almost unavoidable. Responding to the rise in both the influence of judges and the competitiveness of their seats over the last three decades,[19] the cost of judicial election campaigns skyrocketed, both in terms of fundraising and spending.[20] Whereas in 1990 the average judge running in partisan supreme court races spent $425,000 on her campaign (an amount that was already drastically higher than it was just a few election cycles earlier),[21] judges were spending $1.5 million in races just fourteen years later.[22] Judicial candidates in non-partisan races saw a similar spike in spending during that period, moving from $300,000 to $600,000.[23] Data on lower-level courts is less widely available but generally show a similarly seismic shift in the scale of campaigns leading up to the turn of the century.[24]

---

to select intermediate judges, and 22 using them to select high-court judges. Of the 17 states that do not hold popular elections for the initial selection of judges, 6 hold retention elections, meaning that the number of states in which there are no judicial elections at all currently sits at 11. (National Center for State Courts, *Methods of Judicial Selection* (2018), available at http://www.judicialselection.com/judicial_selection/methods/selection_of_judges.cfm?state=).

[19] Bonneau documents that just from 1990 to 2000, the percent of judicial elections for state supreme court seats that were contested rose from 68% to 95% and from 44% to 75% for partisan and non-partisan elections, respectively. Chris W. Bonneau, *Patterns of Campaign Spending and Electoral Competition in State Supreme Court Elections*, 25 JUST. SYS. J. 21, 27 at tbl.6 (2004).

[20] *See* Chris W. Bonneau, *What Price Justice(s)? Understanding Campaign Spending in State Supreme Court Elections*, STATE POL. & POL'Y Q. 5: 107-125 (2005).

[21] *See* American Bar Association, *Report and Recommendations of the Task Force on Lawyers' Contributions to Judicial Campaigns* (1988).

[22] Chris W. Bonneau, *The Dynamics of Campaign Spending in State Supreme Court Elections*, in RUNNING FOR JUDGE: THE RISING POLITICAL, FINANCIAL, AND LEGAL STAKES OF JUDICIAL ELECTIONS 63 fig.4.1 (Matthew J. Streb ed., 2007).

[23] Chris W. Bonneau, *The Dynamics of Campaign Spending in State Supreme Court Elections*, in RUNNING FOR JUDGE: THE RISING POLITICAL, FINANCIAL, AND LEGAL STAKES OF JUDICIAL ELECTIONS 63 fig.4.1 (Matthew J. Streb ed., 2007).

[24] *See, e.g.*, Phillip L. Dubois, *Voting Cues in Nonpartisan Trial Court Elections: A Multivariate Assessment*, 18 L. & SOC. REV. 395 (1984) (reporting that the cost of election in the California Superior Courts (the trial courts) doubled from 1978 to 1982); California Commission on Campaign Financing, *The Price of Justice: A Los Angeles Area Case Study in Judicial Campaign Financing* (1995) (reporting that the "spending for Los Angeles County Superior Court races increased 22-fold" (at 51) between 1976 and 1994).

More recently, the cost of these elections has stabilized and even dipped in some states,[25] but the days in which judicial elections were "low-key affairs"[26] are clearly over.

To complicate matters, studies also show that a significant percentage of donations to judicial campaigns—in some cases a majority of them—come from potential participants in that judge's courtroom.[27] The most recent reports by the Brennan Center calculate that 31.7 percent of the over $40 million raised by judicial candidates from 2015 to 2016 come from lawyers and lobbyists, while an additional 24.1 percent come from business interests.[28]

i. Effect on Judicial Behavior: The obvious question stemming from the recent trends in judicial campaign financing is whether this money is potentially influencing the judges' legal decisions.[29] Many—primarily judges—deny that such a connection exists. In the face of polls that indicated that two-thirds of his electorate questioned his ability to be impartial in an embarrassingly public Supreme Court case, West Virginia Supreme Court Justice Benjamin vehemently denied that the over $3 million in independent spending on his behalf had anything

---

[25] *See* Chris W. Bonneau, *Fundraising and Spending in State Supreme Court Elections*, at 85 tbl. 5.1, in JUDICIAL ELECTIONS IN THE 21ST CENTURY (Chris W. Bonneau & Melinda Gann Hall eds., 2017); Brian Frederick & Matthew J. Streb, *Paying the Price for a Seat on the Bench: Campaign Spending in Contested State Intermediate Appellate Court Elections*, 8 St. POL. & POL'Y. Q. 410 (2008).

[26] Peter D. Webster, *Selection and Retention of Judges: Is There One "Best" Method?*, 23 FLA. ST. U. L. REV. 1, at 19 (1995).

[27] *See, e.g.*, Paul D. Carrington & Adam R. Long, *The Independence and Democratic Accountability of the Supreme Court of Ohio*, 30 CAP. U. L. REV. 455, at 474 (2002) (positing that "[o]ften, lawyers or litigants who are likely to appear before the judge constitute large proportions of the contributions to judicial candidates."); Shira J. Goodman et al., *What's More Important: Electing Judges or Judicial Independence? It's Time for Pennsylvania to Choose Judicial Independence*, 48 DUQ. L. REV. 859, at 865 (2010) (finding that in over 60 percent of the cases that were heard by the Pennsylvania Supreme Court featured a donor as either a party or attorney). See also Siefert v. Alexander, 608 F.3d at 990 (noting that lawyers who "appear before the candidate who wins" are often the key contributors to judicial campaigns).

[28] Alicia Bannon, Cathleen List & Peter Hardin, *Who Pays for Judicial Races?: The Politics of Judicial Elections 2015-2016*. Brennan Center for Justice and National Institute on Money and State Politics, at 11 (2017).

[29] Another key concern is whether the system puts undue pressure on attorneys to donate to judges. *See, e.g.*, Mary Flood, *Got Money? I Beat the Guy You Supported*, Chron (Nov. 19, 2008 at 1:34 PM), https://blog.chron.com/legaltrade/2008/11/got-money-i-beat-the-guy-you-supported/ (reporting on an email sent to a Texas attorney from Court of Appeals Judge Jim Sharp which read in part, "I trust that you will see your way clear to contribute to my campaign account in an amount reflective of the $2,000 contribution you made towards my defeat…").

to do with his decision in *Caperton*.[30] Others believe that the connection between contributions and judicial decisionmaking is obvious, particularly when a contributor is a participant or has stake in the outcome of a case;[31] to not feel the need to engage in reciprocity would "would defy bedrock social norms."[32] Even judges who deny the influence that money may have on their decisions bemoan the position that fundraising puts them in. As one judge colorfully put it: "I never felt so much like a hooker down by the bus station in any race I've ever been in as I did in a judicial race. Everyone interested in contributing has very specific interests. They mean to be buying a vote."[33]

But providing evidence of those effects is difficult, an empirical reality that has been well established by scholars attempting to identify the causal impact of campaign donations on the behavior of legislators and executives.[34] Although it has been repeatedly shown that campaign donors (attorneys, parties, and special interest groups) fare better in the courtroom than non-donors,[35] such observations are consistent with at least two explanations. On the one hand,

---

[30] Caperton v. A.T. Massey Coal Co. 129 S. Ct. 2252 (2009).

[31] See Thomas M. Susman, *Reciprocity, Denial, and the Appearance of Impropriety: Why Self-Recusal Cannot Remedy the Influence of Campaign Contributions on Judges' Decisions*, 26 J. L. & POL. 359, 366 (2011) (for a discussion of the "reciprocity principle). See also the recognition of this problem in *Caperton:* "Though not a bribe or criminal influence, Justice Benjamin would nevertheless feel a debt of gratitude to Blankenship for his extraordinary efforts to get him elected." (556 U.S. at 882).

[32] Dmitry Bam, *Recusal Failure*, 18 LEG. & PUB. POL'Y 631, at 639 (2015).

[33] Adam Liptak & Janet Roberts, *Campaign Cash Mirrors a High Court's Rulings*, N.Y. TIMES, Oct. 1, 2006, § 1, at 11, quoting Justice Paul E. Pfeifer (a member of the Ohio Supreme Court).

[34] For an early review of this problem, *see* Henry W. Chappell, *Campaign Contributions and Congressional Voting: A Simultaneous Probit-Tobit Model*, 64 Rev. Econ. & Stat. 77 (1982). *See also* Stacy Gordon, *All Votes Are Not Created Equal: Campaign Contributions and Critical Votes,* 63 J. Pol. 249 (2003).

[35] See, e.g., Stephen J. Ware, *Money, Politics, and Judicial Decisions: A Case Study of Arbitration Law in Alabama*, J. L. & POL. 25: 645 (1999) (finding that an individual judge's propensity for supporting arbitration is strongly correlated with the source of that judge's election funding, with judges receiving funding from businesses more likely to support arbitration than those receiving funding from plaintiff's lawyers); Eric N. Waltenburg & Charles S. Lopeman, Tort Decisions and Campaign Dollars, 28 SE POL. REV. 241 (2000); Vernon Valentine Palmer & John Levendis, T*he Louisiana Supreme Court in Question: An Empirical and Statistical Study of the Effect of Campaign Money on the Judicial Function*, 82 TULANE L. REV. 1291 (2008); Michael S. Kang & Joanna M. Shepherd, *The Partisan Price of Justice: An Empirical Analysis of Campaign Contributions and Judicial Decisions*, 86 N. Y. U. L. REV. 69 (2011); Eric N. Waltenburg & Charles S. Lopeman, *2000 Tort Decisions and Campaign Dollars*, 28 SE POL. REV. 241,

judges may be favoring contributors in their rulings (either intentionally or unintentionally—more on this distinction in *infra* Subpart III.A), leading to better outcomes for cases in which those contributors are participating. On the other hand, contributors may simply be donating to judges who harbor a legal or political sensibility that will naturally (and innocently) lead them to vote in ways that benefit the contributor.[36] This causal question is made all the more difficult by the fact that it does not easily lend itself to experimentation—the methodological approach that best addresses causal concerns.[37]

Nonetheless, recent studies have used clever empirical methods that—to some extent—get around some of the causal problems inherent in the enterprise. For example, a pair of studies instrumented for campaign contributions using the presence of a public defender and the incumbency status of the judge during the campaign—variables the authors argued were exogenous predictors of the amount of campaign contributions a judge receives.[38] Another approach has been to identify existing institutional features that might create discrepancies in contributions across otherwise similar judges. A recent study by Hazelton et al. utilized North

248, 256 (2000) (examining tort cases before supreme courts in Alabama, Kentucky, and Ohio, finding that judges rule in line with the interests of donors when they are approaching reelection).

[36] See Chris W. Bonneau, *A Survey of Empirical Evidence Concerning Judicial Elections,* Federalist Society White Paper (March) (providing a particularly good explanation of this two-way causal problem); and Douglas D. Roscoe & Shannon Jenkins, *A Meta-Analysis of Campaign Contributions' Impact on Roll Call Voting,* SOC. SCI. Q. 86:52–68 (2005) (discussing this sort of "friendly giving" in the context of the legislative branch).

[37] This does not mean that experimentation in this arena is not possible. For a particularly fantastic example of how field experiments can be applied to seemingly implausible circumstances, see Joshua L. Kalla and David Broockman, *Campaign Contributions Facilitate Access to Congressional Officials: A Randomized Field Experiment*, 60 AM. J. POL. SCI. 545 (2016). The researchers partnered with a political organization whose members had made donations to over 191 congressional representatives and randomly varied whether it revealed the contributions when attempting to schedule meetings with the representatives. Kalla and Broockman found that the congressional offices were over three times willing to meet with donors than with non-donors.

[38] Chris W. Bonneau, *The Effects of Campaign Spending in State Supreme Court Elections*, 60 POL. RES. Q. 489 (2007); Chris W. Bonneau & Damon M. Cann, *Campaign Contributions, Judicial Decisionmaking, and Institutional Context*, Unpublished Manuscript (2009).

This approach, while novel, has been criticized based on the use of instruments that are likely not independent of the error term. *See, e.g.*, Herbert M. Kritzer, *Impact of Judicial Elections on Judicial Decisions*, 12 ANN. REV. L. SOC. SCI. 353 (2016).

Carolina's opt-in public financing system for supreme court candidates and compared the decisions of judges before and after they entered into the system against those who privately funded all of their campaigns.[39] The authors found that the advantage that attorney donors had when all judges ran privately funded campaigns diminished when they shifted to public funding, reasonably strong evidence that the donations were affecting decisions.[40] Similarly, a number of nationwide studies by Shepherd compared state supreme court judges who were potentially running for an additional term against those who were prevented from doing so due to mandatory retirement rules.[41] Shepherd's studies focus primarily on the relationship between case outcomes and the business interests of donors and show a strong relationship between the two. These findings were recently validated and extended by another national study conducted by Heise, who found that non-business interests of donors are also influential on judicial behavior.[42]

The integrated implications of these studies are nuanced, in large part because the influence of campaign money on case outcomes seemed to depend on the circumstances of the election and the nature of the campaign contribution. Not surprisingly, judges who won by slimmer margins in their previous election were more likely to rule in favor of contributors than judges who won by wider margins,[43] and judges appeared to more greatly favor the interests of

[39] Morgan L. W. Hazelton, Jacob M. Montgomery, & Brendan Nyhan*, Does Public Financing Affect Judicial Behavior? Evidence from the North Carolina Supreme Court*, 44 AM. POL. RES. 587 (2015).

[40] The key causal concern here is that the judges who opted into the publicly funded systems might be (and probably are) systematically different than the judges who did not.

[41] Joanna M. Shepherd, *Money, Politics, and Impartial Justice*, 58 DUKE L. J. 623 (2009); Michael S. Kang & Joanna M. Shepherd, *The Partisan Price of Justice: An Empirical Analysis of Campaign Contributions and Judicial Decisions*, 86 N.Y.U. L. REV. 69 (2011).

[42] Michael Heise, *The Complicated Business of State Supreme Court Elections: An Empirical Perspective*, Cornell Legal Studies Research Paper (2018).

[43] *See* Ryan J. Rebe, *Analyzing the Link Between Dollars and Decisions: A Multi-state Study of Campaign Contributions and Judicial Decision Making*, 35 AM. REV. POL. 65 (2016).

their donors when approaching reelection.[44] The structure of the electoral system seems to be a factor as well: Bonneau and Cann found that judges selected in the partisan races of Michigan and Texas appeared to be more influenced by contributions than judges in the non-partisan elections of Nevada, although those differences may also have been due to other unobserved variances in the individual states' systems or politics.[45]

The weight of the current evidence suggests a positive causal relationship between the source of donations and the outcomes of court cases, with some studies showing that donation differences of just $100 can have a significant effect on how the judge will determine the outcome of cases.[46] Nonetheless, there are still gaps in this relatively new body of scholarship. Of particular importance in the context of this Chapter's empirical study, these studies have almost always analyzed the relationship between money and judicial decisions in state *supreme* courts, giving little attention to intermediate and trial courts (a bias that is present throughout the entire subfield of courts and judicial studies).[47]

ii. Effect on Public Perception: Even if one is skeptical about the causal link between political donations and the outcomes of individual cases, donations to judges can still be detrimental to the legal system if there is a perceived relationship between the two. In fact, some evidence suggests that a public perception of impropriety and partiality is even more concerning

---

[44] *See* Eric N. Waltenburg & Charles S. Lopeman, *2000 Tort Decisions and Campaign Dollars*, 28 SE. POL. REV. SE. POL. REV. 241 (2000) (examining tort cases before supreme courts in Alabama, Kentucky, and Ohio across the judges' election terms).

[45] Chris W. Bonneau & Damon M. Cann, *Campaign Contributions, Judicial Decisionmaking, and Institutional Context*, Unpublished Manuscript (2009).

[46] *See* Damon M. Cann, Chris W. Bonneau, & Brent D. Boyea, *Campaign Contributions and Judicial Decisions in Partisan and Nonpartisan Elections*, in NEW DIRECTIONS IN JUDICIAL POLITICS (Kevin T. McGuire ed. 2012). After controlling for ideology and case-specific variables, the authors find that there is a significant relationship between the donations of attorneys and the outcomes of cases. "For a judge who received $100 more from the liberal party's attorney(s), the odds of a liberal decision are more than double the odds (specifically 2.29 times greater) where the respective parties' contributions are at parity. For a judge who receives $100 more contributions from the conservative party, the odds of a liberal decision decrease by a factor of 0.44" (at 49).

[47] *See, e.g.*, Jeffrey J. Rachlinski & Andrew J. Wistrich, *Judging the Judiciary by the Numbers: Empirical Research on Judges*, 13 ANN. R. L. & SOC. SCI. 1, at 3-4 (2017) (noting that "most research on judges emphasizes decisions of the US Supreme Court…").

than actual bias.[48] And while the evidence supporting the link between campaign spending and judicial behavior is open to criticism, the evidence that the public, lawyers, and even judges are concerned with the effect that donations have on judicial decisionmaking is undeniably strong.

Polling data has consistently shown that nearly every important demographic group sees the increased role of money in judicial elections as both a source of bias in individual cases and a general threat to the legitimacy of the court system. 76 percent of voters in in a national poll believed that campaign contributions had at least "some influence" on judicial decisions, and an additional 14 percent believed that there is "a little influence" on case outcomes.[49] In another poll, more than 90 percent of those polled believed that judges should not hear cases involving contributors.[50] Similarly, 90 percent of business leaders are concerned about the role that money plays in judicial behavior.[51]

Even lawyers and judges, presumably the individuals who would be best informed regarding the true impact that donations play in court decisionmaking, appear to be concerned about the impact that campaign contributions have on case outcomes. 46 percent of state judges polled in 2001 indicated that they thought campaign contributions influenced judges' decisions, compared to 36 percent who believed there was no influence,[52] and 60 percent of these judges supported proposals to make recusal mandatory in cases that feature parties who had financially supported the presiding judge's campaign.[53]

---

[48] *See, e.g.*, Tom Tyler, *Social Justice: Outcome and Procedure*, 35 INT. J. PSY. 117 (2010).

[49] Justice at Stake, *Justice at Stake Frequency Questionnaire,* available at http://www.justiceatstake.org/media/cms/ JASNationalSurveyResults_6F537F99272D4.pdf.

[50] Justice at Stake, *Polls of Voters, Business Leaders and Judges Themselves Shows Overwhelming Concern About the Impact of Special Interest Money on the Integrity of American Courts*, available at http://www.justiceatstake.org/ resources/polls.cfm.

[51] Christian W. Peck, *Attitudes and Views of American Business Leaders on State Judicial Elections and Political Contributions to Judges*, Zogby International, (2007), available at http://www.justiceatstake.org /media/cms/CED_FINAL_ repor_ons_14MAY07_BED4DF4955B01.pdf

[52] 16 percent said they did not know whether contributions influenced decisions, and 2 percent did not answer the question. Justice at Stake*, State Judges Frequency Questionnaire* (2002), available at http://www.justiceatstake.org /media/cms/JASJudgesSurveyResults_EA8838C0504A5.pdf.

[53] Justice at Stake*, State Judges Frequency Questionnaire* (2002), available at

A growing literature of survey experiments led by the work of Gibson and Caldeira have further explored the mechanisms behind the public's views on campaign spending in judicial elections.[54] By randomly varying the content of vignettes about a candidate's judicial campaign—including the existence, type (independent or direct), and size of campaign contributions; whether the judge accepted the contributions; and whether the judge recused from cases that featured campaign contributors—the authors are able to identify the extent to which each of those factors impact perceptions of judicial legitimacy. The results of their nationwide study show that when judicial candidates accept contributions, survey participants are more than 30 percentage points less likely to believe that the judge can be fair and impartial. Perceptions of legitimacy are also damaged by judges who hear cases that feature campaign contributors relative to those who recuse (a 20 percentage-point difference).

C. Recusal and Judicial Campaign Finance

i. Caperton v. Massey Coal: The touchstone case for discussing the confluence of judicial politics and recusal is undoubtedly *Caperton v. A.T. Massey Coal Co.*[55] It features prominently in the introductions of nearly all post-2009 law review articles that cover judicial campaign finance or recusal, and it is easy to see why: it is a spectacular example of the potential threats to judicial impartiality and legitimacy that come from judicial elections. But the extreme circumstances of the case also make it a somewhat unhelpful case study for understanding the current status of recusal law in the United States. Although the Supreme Court ultimately found the judge's failure to recuse to be a violation of due process, the resulting precedent is so ambiguous and tied to such extraordinary facts, that states were given little guidance in how to

http://www.justiceatstake.org /media/cms/JASJudgesSurveyResults_EA8838C0504A5.pdf.

[54] James L. Gibson & Gregory A. Caldeira, *Judicial Impartiality, Campaign Contributions, and Recusals: Results from a National Survey*, 10 J. EMP. L. STUD. 76 (2013); James L. Gibson & Gregory A. Caldeira, *Campaign Support, Conflicts of Interest, and Judicial Impartiality: Can Recusals Rescue the Legitimacy of Courts?,* 74 J. POL. 18 (2012).

[55] Caperton v. A.T. Massey Coal Co., 129 S. Ct. 2252 (2009).

address the frequent but much more mundane conflicts created by campaign donors in the courtroom. Nonetheless, the expressed confidence of the judge in question in being able to remain unbiased and the contrasting perspectives of the Supreme Court justices in the *Caperton* opinions are emblematic of the diversity of views regarding the nature of judicial campaign finance and the extent to which we are willing to give judges a presumption of impartiality. Consequentially, we feel the case warrants an overview (plus, the facts truly *are* fascinating).

Caperton originated as a contractual suit in the West Virginia trial court in which Hugh Caperton and several other plaintiffs laid suit against the A.T. Massey Coal Company for a contractual dispute.[56] The trial court awarded $50 million in compensatory and punitive damages to the plaintiffs.

Knowing that the case would be appealed to the West Virginia Supreme Court,[57] Massey's chief executive officer and president, Benjamin Blankenship, threw substantial financial support behind Brent Benjamin, a local attorney who was challenging Justice McGraw, an incumbent in the 2004 state supreme court elections. Blankenship spent over $500,000 in independent expenditures on political advertisements and direct mailers supporting Benjamin and donated just under $2.5 million to "And For The Sake Of The Kids," a section-527 organization supporting Benjamin's run for office. In total, Blankenship spent roughly $3 million on Benjamin's campaign, an amount that Caperton alleged was $1 million more than what Benjamin's and McGraw's campaign committees spent combined.

In November 2004, Benjamin defeated McGraw to become the first Republican Supreme Court justice in West Virginia in over 80 years, and Benjamin took office in January 2005. In October 2005, Caperton preemptively moved for Benjamin's removal based on West Virginia's Code of Judicial Conduct and the Due Process Clause of the U.S. Constitution. After reviewing

---

[56] Caperton v. A.T. Massey Coal Co., 129 S. Ct. 2252, 2 (2009).

[57] West Virginia has a two-court system for substantial cases such as *Caperton*. *See West Virginia State Court Structure Chart*, Court Statistics Project, available at *http://www.courtstatistics.org/Other-Pages/State_Court_Structure_Charts/West-Virginia.aspx*

the motion, Benjamin found "no objective information . . . to show that [he] has a bias for or against any litigant, that [he] has prejudged the matters which comprise this litigation, or that [he] will be anything but fair and impartial."[58] Massey subsequently filed petition for appeal in December 2006, and the West Virginia Supreme Court granted review. In a 3-2 decision, the court reversed the trial court's decision in favor of Massey Coal based on a forum-selection clause in one of the relevant contracts that barred the suit from review in West Virginia and based on res judicata stemming from a previous judgement to which Massey Coal was not a party.

Caperton subsequently sought rehearing based in part on recent evidence that had emerged showing that Justice Maynard—one of the three votes in favor of Massey Coal—had vacationed with Blankenship in the French Rivera while the case was pending in the West Virginia Supreme Court.[59] Caperton submitted a motion to recuse to Maynard and again submitted a motion to recuse to Benjamin for the campaign support of Blankenship. At the same time, Massey Coal sought recusal of Justice Starcher—one of the two dissenting votes—due to his public criticism of Benjamin's failure to recuse before the original hearing. Maynard and Starcher recused, but Benjamin denied the recusal motion against him.

The court, led by Acting-Chief Justice Benjamin, and composed of three of the original justices and two new justices selected by Benjamin, reheard the case, but only after a third recusal request against Benjamin was denied. With a recent push poll showing that over two-thirds of West Virginians questioned Benjamin's ability to be fair and impartial, Caperton argued that, in accordance with West Virginia's Code of Judicial Conduct, it was clear that "a reasonable and prudent person" would question Benjamin's ability to be impartial. In April 2008, the five-person panel again found in favor of Massey Coal in a 3-2 decision.

---

[58] Caperton v. A.T. Massey Coal Co., 129 S. Ct. 2252, 3 (2009).

[59] *See* Adam Liptak, *Motion Ties W. Virginia Justice to Coal Executive,* N.Y. TIMES, Jan. 15, 2008, available at https://www.nytimes.com/2008/01/15/us/15court.html.

The U.S. Supreme Court granted cert in the case, and in a 5-4 decision found in favor of Caperton, overturning the West Virginia Supreme Court's ruling. The U.S. Supreme Court's previous standard required recusal when the probability of judicial bias "is too high to be constitutionally tolerable,"[60] and the Court found that the probability of bias in this case reached that level.

The constitutional question at issue, at least as it was framed by the Kennedy-penned majority opinion, was specifically whether the donations and support given to Benjamin by Blankenship (the validity of which were not in question) created an intolerable risk of bias, not "whether in fact [the justice] was influenced."[61] In fact, the majority made it clear that they were not questioning Benjamin's impartiality and that due process "may sometimes bar trial by judges who have no actual bias and who would do their very best to weigh the scales of justice equally between contending parties."[62]

The dissenting justices disagreed with the majority's approach in a number of ways that are worth highlighting. First, they noted that the Court's previous jurisprudence had never found a failure to recuse to be a violation of due process except when the judge has a financial interest in the case or when the judge is deciding on a criminal contempt resulting from previous hostility to that same judge. The Court had explicitly stated that even issues as problematic as kinship to the judge may not be constitutional in nature.[63] Second, in addition to venturing beyond previous rulings, they felt that the majority had based its decision on a standard—probability of bias—that "fail[ed] to provide clear, workable guidance for future cases."[64] Third, the dissenting justices held a much stronger presumption of impartiality in regard to judges.

---

[60] Withrow v. Larkin, 421 U.S. 35, at 47 (1975).

[61] Caperton v. A.T. Massey Coal Co., 129 S. Ct. 2252, 8 (2009), quoting Aetna Life Ins. Co. v. Lavoie, 475 U. S. 813, 825 (1986).

[62] In re Murchison, 349 U.S. 133, 136 (1955).

[63] *Tumey v. Ohio*, 273 U.S. 510, 523 (1927); see also Aetna Life Ins. Co. v. Lavoie, 475 U. S. 813, 820 (1986).

[64] Caperton v. A.T. Massey Coal Co., 129 S. Ct. 2252 (2009) (Roberts, J., dissenting).

They questioned whether the facts of the cases were really all that extreme, emphasizing that the vast majority of Blankenship's support was made via independent expenditures and was therefore out of Benjamin's control.[65] Scalia's dissent even argued that the majority's ruling was more deleterious to the public's perception of the judiciary than the facts of the case themselves.[66]

In some ways, this ruling was seen by many as a monumental shift in the approach that the Supreme Court takes toward elections and campaign spending—it was the first time that the Court had ever found a failure to recuse due to campaign contributions to be a violation of due process, and it was one of the few times where the Court was willing to push back against the growing influence of campaign spending in elections. In regard to recusal jurisprudence, however, the decision provided very little guidance. The Court made clear that there exists some "constitutional floor"[67] at which point judges must recuse due to campaign contributions but left questions regarding the possible bias associated with smaller, more common, donations (that is, donations less than $3 million) as "matters merely of legislative discretion."[68]

ii. Recusal for Campaign Contributions: Despite the Supreme Court punting on the specifics of recusal in relation to campaign contributions, recusal is still understood by many as the most practical antidote to judicial impartiality. Indeed, in the case of campaign contributions, some have argued that recusal is not only the best means for dealing with bias but "the *only* effective means to ensure the impartiality of elected judges."[69] Dmitry Bam, one of the leading scholars on recusal policy (and a skeptic regarding its efficacy in regard to campaign

---

[65] Caperton v. A.T. Massey Coal Co., 129 S. Ct. 2252 (2009) (Roberts, J., dissenting at 13).

[66] Caperton v. A.T. Massey Coal Co., 129 S. Ct. 2252 (2009) (Scalia, J., dissenting at 1).

[67] Caperton v. A.T. Massey Coal Co., 129 S. Ct. 2252, 19 (2009), quoting Bracy v. Gramley, 520 U. S. 899, 904 (1997).

[68] Caperton v. A.T. Massey Coal Co., 129 S. Ct. 2252, 6 (2009), quoting Tumey v. Ohio, 273 U. S. 510, at 523 (1927).

[69] Molly McLucas, *The Need for Effective Recusal Standards for an Elected Judiciary*, 42 LOY. LA. L. REV. 671, 692 (2009) (emphasis added).

contributions), explains: "[r]ecusal has tremendous allure because, in theory, it allows us to ensure judicial impartiality at the point of delivery."[70] In other words, it provides a post hoc remedy to a problem that is an inherent part of the current political regime without a wholesale dismantling of the regime itself.

However, the particular legal approach that states should take to regulating recusal in relation to campaign contributors is still an open question. Much of this discussion has been led by the American Bar Association and their Model Code of Judicial Conduct, which has no force of law but is widely seen as the preeminent source for judicial standards in this area. The ABA first suggested recusal as a solution for contributor bias in a 1999 amendment, and the current rule (most recently amended in 2007) remains largely the same.[71] Rule 2.11 calls for automatic recusal when "The judge knows or learns by means of a timely motion that a party, a party's lawyer, or the law firm of a party's lawyer has within the previous [insert number] year[s] made aggregate contributions to the judge's campaign in an amount that is greater than $[insert amount] for an individual or $[insert amount] for an entity [is reasonable and appropriate for an individual or an entity]," leaving the specific contribution amounts at which the rule is triggered up to the individual states. This provision has subsequently led to a complex (and often heated) debate within the ABA's organizational bodies, both regarding the ethical implications it imposes for judges and the potential it creates for attorneys and parties to intentionally avoid certain judges through making strategic contributions (see a discussion of this "pay-not-to-play" problem in *infra* Part VI of this Chapter).

Despite widespread adoption of the ABA Model Code as a basis for their state codes of judicial conduct, only a handful of states have implemented Rule 2.11(4)'s inclusion of campaign contributions in the list of enumerated *per se* recusal requirements.[72] Most prominently,

---

[70] Dmitry Bam, *Recusal Failure*, 18 LEG. & PUB. POL'Y 631 (2015).

[71] Model Code of Judicial Conduct Canon 3E(1)(e) (1999) (Am. Bar Ass'n, amended 2011).

[72] See American Bar Association, *Comparison of Jurisdictional Codes of Judicial Conduct to Model Code of Judicial Conduct*, available at

California passed a 2010 amendment to their Code of Civil Procedure to require recusal for trial judges if a party or an attorney in a case had donated more than $1,500 in the previous six years[73] or for any donation amount if it a reasonable person might "entertain a doubt that the judge would be able to be impartial."[74] The California Supreme Court also amended the code of judicial ethics to require recusal for appellate judges for donations above $5,000.[75] Similarly, Arizona implemented mandatory recusal for aggregate contributions that exceed the campaign contribution limits at the time,[76] and in 2014 Alabama adopted rules stipulating that judges must recuse if the contributions of a lawyer or party in a case exceed ten, fifteen, or twenty-five percent of the total money raised by the appellate, circuit, and district court judge, respectively.[77] On a much more extreme level, Utah similarly requires recusal for donations, except the cap is any aggregate contribution amount above $50.[78]

Interestingly, two state supreme courts—one of which is a venue for this Chapter's experiment—quashed similar reforms for *per se* rules. In 2010, both the League of Women

---

https://www.americanbar.org/groups/professional_responsibility/resources/judicial_ethics_regulation/comparison.html.

[73] California Code of Civil Procedure § 170.1(9)(A).

[74] California Code of Civil Procedure § 170.1(9)(A) Comments.

[75] California Supreme Court Code of Judicial Ethics Canon 3E(5)(j), which reads in whole: "The justice has received a campaign contribution of $5,000 or more from a party or lawyer in a matter that is before the court, and either of the following applies: (i) The contribution was received in support of the justice's last election, if the last election was within the last six years; or (ii) The contribution was received in anticipation of an upcoming election. Notwithstanding Canon 3E(5)(j), a justice shall be disqualified based on a contribution of a lesser amount if required by Canon 3E(4). The disqualification required under Canon 3E(5)(j) may be waived if all parties that did not make the contribution agree to waive the disqualification."

[76] Arizona's contribution limits change systematically on a yearly basis according to A.R.S. Section 16-905.

[77] Alabama Laws Act 2014-455 http://legiscan.com/AL/text/HB543/2014. Interestingly, Alabama had previously had a mandatory recusal cap at $4,000, that was never implemented. See Little v. Strange, 796 F. Supp. 2d 1314 (2011) (in which the U.S. District Court for the Middle District of Alabama dismissed a suit challenging the statute because of a lack of ripeness).

[78] Utah Supreme Court Code Rule 2011(A)(4), requiring disqualification when ""the judge knows or learns by means of a timely motion that a party, a party's lawyer, or the law firm of a party's lawyer has within the previous three years made aggregate contributions to the judge's retention in an amount that is greater than $50." Note that one reason why the cap is so low is because Utah judges fundraise only during retention elections, which are considerably less expensive than competitive races.

Voters of Wisconsin[79] and a former Wisconsin Supreme Court Justice[80] submitted petitions to require recusal for donations above $1,000 and $10,000. Responding instead to two contrary proposals filed by the Wisconsin Realtors Association[81] and Wisconsin Manufacturers and Commerce,[82] the court included an explicit provision in the supreme court rules that clarified that previous endorsements, contributions, or independent expenditures by participants in a case do not induce automatic removal.[83] Similarly, Nevada's supreme court rejected a provision recommended by the Commission on the Amendment to the Nevada Code of Judicial Conduct that required recusal for contributions exceeding $50,000 or five percent of the judge's total fundraising.

Other states have explicitly incorporated considerations for campaign contributions into their recusal regimes but have stopped short of requiring the judge to remove herself for certain amounts. Mississippi now allows parties to file motions to recuse[84] based on donations of more than $2,000 for appellate judges and $1,000 for trial judges,[85] but this requires only that the motion be "considered and subject to appellate review as provided for other motions for recusal."[86] Georgia incorporated a list of factors that "may be considered" to determine

---

[79] Petition of The League of Women Voters of Wisconsin Education Fund, *In re creation of rules for when a party or a lawyer in a case made contribution affecting a judicial campaign,* Wisconsin Supreme Court, available at https://wicourts.gov/supreme/docs/0816petitionamend.pdf.

[80] Petition of Justice William A. Bablitch, *In the Matter of Amending the Code of Judicial Conduct,* Wisconsin Supreme Court, available at https://www.wicourts.gov/supreme/docs/0911petition.pdf.

[81] Petition of The Wisconsin Realtors Association, *In the Matter of Amending the Rules of Judicial Conduct,* Wisconsin Supreme Court, available at https://wicourts.gov/supreme/docs/0825petition.pdf.

[82] Petition of Wisconsin Manufacturers & Commerce, *In the Matter of Amending the Rules of Judicial Conduct,* Wisconsin Supreme Court, available at https://www.wicourts.gov/supreme/docs/0910petition.pdf.

[83] Wisconsin Supreme Court Rules 60.04 (7) (which discusses direct contributions) and 60.04 (8) (which discusses independent expenditures). See also, Tom Solberg, *Supreme Court Revises Judicial Recusal Guidelines,* State Bar of Wisconsin, July 9, 2010, available at https://www.wisbar.org/NewsPublications/InsideTrack/Pages/Article.aspx?Volume=2&Issue=37&ArticleID= 5912

[84] This is distinct from a requirement of automatic recusal.

[85] Mississippi Code of Judicial Conduct, Canon 3E(2).

[86] Mississippi Code of Judicial Conduct, Canon 3E(2).

impartiality due to campaign support (which includes donation amount in addition to the timing and impact of that donation), but clarifies that "[t]here is a rebuttable presumption that there is no per se basis for disqualification where the aggregate contributions are equal to or less than the maximum allowable contribution permitted by law."[87] Since 2010, similar provisions or amendments to existing code that address campaign contributions have also been made in Iowa,[88] Michigan,[89] Oklahoma,[90] Pennsylvania,[91] Tennessee,[92] and Washington;[93] and Arkansas,[94] Missouri,[95] New Mexico,[96] and North Dakota,[97] have added relevant language in the comments to their rules.[98]

For these states with recusal laws and rules that simply mention campaign contributions as a potential source of bias or do not address such relationships at all, whether or not a judge should recuse from cases in which a donor is a participant is governed by the generalized bias or

---

[87] Georgia Code of Judicial Conduct, Canon 3E(1)(d)).

[88] Iowa Code of Judicial Conduct, Chapter 51:2.11(A)(4).

[89] Michigan Rules of Court, Rule 2.003(C)(1)(b).

[90] Oklahoma Code of Judicial Conduct, Rule 2.11A(4).

[91] 207 Pennsylvania Code Rule 2.11A(4). See also Robert A. Graci, Press Release: Board Issues Statements of Policy Regarding Investigations of Campaign Contributions and Electronic Communications, Commonwealth of Pennsylvania Judicial Conduct Board (November 4, 2016), available at http://judicialconductboardofpa.org/wp-content/uploads/11-04-2016-Press-Release-Board-Issues-Statements-of-Policy-Regarding-Investigations-of-Campaign-Contributions-Electronic-Communications.pdf.

[92] Rule 10 Tennessee Code of Judicial Conduct, Cannon 2.11A(4). See also Comment 7.

[93] Washington Code of Judicial Conduct, Cannon 2.11(D).

[94] Arkansas Code of Judicial Conduct Rule 2.11, comment 4 (which reads: "The fact that a lawyer in a proceeding, or a litigant, contributed to the judge's campaign, or publicly supported the judge in his or her election does not of itself disqualify the judge. However, the size of contributions, the degree of involvement in the campaign, the timing of the campaign and the proceeding, the issues involved in the proceeding, and other factors known to the judge may raise questions as to the judge's impartiality under paragraph (A).")

[95] Missouri Supreme Court Rule 2-4.2 Comment.

[96] New Mexico Supreme Court Rule 21.211 Comment.

[97] North Dakota Code of Judicial Conduct, Cannon 2.11 Comment.

[98] For a thorough overview of these changes and all the recent legal moves relating to recusal and campaign contributions, see Cynthia Gray, *Judicial Disqualification Based on Campaign Contributions*, National Center for State Courts Center for Judicial Ethics (2016).

appearance of bias provisions discussed in *supra* Subpart II.C. As a result, it is not surprising to see that, upon review of cases in which judges have refused to recuse,[99] the majority of state courts have found that judges are not required to recuse due purely to campaign donations if those judges have determined that they can remain impartial.[100] This does not preclude judges from recusing themselves regardless of the absence of a precedential requirement to do so, but it does create a system of recusal that is firmly rooted in the subjective evaluation of the judge herself.

---

[99] When evaluating patterns in the case law, it is important to be aware of how the institutional framework of the courts might result in a biased picture of the legal practice and outcomes in that area. With cases dealing with recusal motions, we should expect to find a disproportionate amount of cases in which the initial judge has refused to recuse because it is unlikely (and legally tenuous) for parties to appeal in cases when the judge has either recused herself or granted a motion to recuse. See Sande L. Buhai, *Federal Judicial Disqualification: A Behavioral and Quantitative Analysis*, 90 Or. L. Rev. 69 (2011) (for a discussion of this limitation in the context of recusals in the Federal Courts).

[100] *See* RICHARD E. FLAMM, JUDICIAL DISQUALIFICATION: RECUSAL AND DISQUALIFICATION OF JUDGES 367-386 (3rd ed. 2017) (providing a short review of the current approach to this issue by U.S. courts, including a summary of a good number of relevant cases). *See, e.g., Rocha v. Ahmad*, 662 S.W.2d 77 (Tex. App. 1983) (not requiring recusal despite the fact that two of the three judges had received thousands in contributions from one of the attorneys and had thrown their victory celebrations at the attorneys' offices); Storms v. Action Wisconsin Inc., 2008 WI 110, 314 Wis. 2d 510, 754 N.W.2d 480 (Wis. 2008) (arguing that there is "no case in Wisconsin or elsewhere" where recusal is mandated simply because a participant in the case made a contribution to the judge's political campaign); *Williams v. Viswanathan*, 65 S.W.3d 685 (Tex. App. 2001) (finding that because a reasonable person would certainly know that judges have to raise money in order to run for election, simply having received a donation from a participant in a case is not enough to warrant required recusal); *Braynen v. State*, 895 So. 1169 (Fla. App. 2005) (finding a judge cannot be required to remove herself purely based on a campaign donation, but in a concurring opinion, it was noted that in "spite of our extant precedent, a judge would be well advised to grant recusal under these circumstances," because a reasonable person could very well "understandably fear" impartiality (at 1170)); Whalen v. Murphy, 943 So. 2d 504 (La. App. 2006) (finding that "the evidence simply does not establish actual bias," (at 508) in a case in which the judge vacated a previously granted summary judgement in favor of a party after the opposing party made a campaign contribution); Shepherdson v. Nigro, 5 F. Supp.2d 305, 308-311 (E.D. Pa. 1998) (acknowledging that the state of campaign finance creates an unfortunate situation but that it would be unfair to punish the judge for accepting legal contributions by forcing recusal).

But see Gude v. State, 289 Ga. 46 709 S.E.2d 206 (Ga. 2011) (not requiring recusal but opining that judges should recuse from cases "involving a party who has previously made an exceptionally large campaign contribution…" (at 50)); Reems v. St. Joseph's Hospital, 536 N.W.2d 666 (N.D. 1995) (not requiring recusal in a case in which one of the lawyers was a co-chair for the judge's campaign finance committee, but noting that campaign contributions are not irrelevant in determining whether recusal is necessary).

III. Why Judges Don't Recuse Themselves and Attorneys Don't Ask Them To

We are not the first to claim that the current system of recusal in the United States is inefficacious. Coming after the controversy surrounding the *Caperton* case, scholarship skeptical of recusal in the context of judicial elections has become a cottage industry.[101] Across this literature, the message—recusal does not work—is the same but the concerns are varied. Below, we attempt consolidate these concerns and build an informal model of judicial behavior that identifies the various factors that disincentivize recusal even when a legitimate conflict of interest exists. This is not a functional formal model—it does not attempt to provide a numerical framework for the decision—but it is an attempt to be comprehensive. As we argue below, the analysis leads to the conclusion that recusal rates, particularly when the recusal question pertains to campaign contributions, are likely very low—even lower than the literature skeptical of recusal might suggest.

The model, which will later be adjusted to explore the incentives surrounding an attorney requesting recusal, focuses on the judge's behavior but accounts for three primary actors: the presiding judge for whom the conflict of interest is a potential issue; the "inside" party, who has the questionable relationship with the judge; and the "outside" party, who does not have the relationship with the judge. In this structuring, we will generally treat the parties to a case and their attorneys as single actors with unified preferences, although as we will highlight in *infra* Subpart III.C, there are instances when the long-term interests and subsequent behavior of the attorney diverge from the short-term interests of her client.

It should be noted that for many readers, particularly those trained in U.S. law schools, this analysis may seem to paint an overly critical image of the American judiciary. The traditional legal pedagogy is strongly influenced by a formalist strain of judicial philosophy that

[101] Morrow notes that, as of 2018, there are 682 law review articles that cite *Caperton*. (Kevin Morrow, *To Judge or Not to Judge? Judge Shopping, Recusal & Judicial Defendants,* Unpublished Manuscript (2018).

sees judges as legal umpires who "call balls and strikes"[102] and are able to apply the law without consideration of extra-legal factors.[103] Our analysis espouses a more realist-oriented approach, arguing that a litany of behavioral considerations—for example, a judge's reputation among peers or her chances for reelection—influence a judge's behavior in addition to and in tandem with the relevant legal considerations at hand.[104] But it does not follow from this approach that all judges intentionally spurn the law to accomplish their own ends. One can, as we do in this Chapter, simultaneously adopt the position that judges are susceptible to a wide variety of incentives and biases and still maintain that most judges are sincerely trying their best and are among the most qualified and capable individuals in the legal profession.

A. Why Judges Don't Recuse Themselves

i. Judicial Latitude in Recusal Determinations: The foundational factor in understanding why judges are likely not inclined to recuse is the procedural structure of the recusal process itself. The entire premise of employing a judge to make legal determinations is based on the well-known edict that "no man is allowed to be a judge in his own cause."[105] And yet, with few exceptions, the judge to whom the recusal motion is filed and the judge who rules on the recusal motion is the very individual whose impartiality is being questioned.[106] This practice of self-recusal is both a problem in and of itself and also lays the groundwork for a host of other considerations that influence the judge's decision to recuse or not.

---

[102] Confirmation Hearing on the Nomination of John G. Roberts, Jr. to be Chief Justice of the United States: Hearing Before the S. Comm. on the Judiciary, 109th Cong. 55 (2005).

[103] See BRIAN Z. TAMANAHA, BEYOND THE FORMALIST-REALIST DIVIDE: THE ROLE OF POLITICS IN JUDGING 30 (2009).

[104] John Leubsdorf, *Theories of Judging and Judge Disqualification,* 62 N.Y.U. L. REV. 237, 247 (1987).

[105] THE FEDERALIST No. 10, at 59 (James Madison) (Jacob E. Cooke ed., 1961)). The full quote states that, "No man is allowed to be a judge in his own cause, because his interest would certainly bias his judgment, and, no improbably, corrupt his integrity."

[106] As an example of an exception, Illinois recusal law allows for the substitution of the judge for recusal determinations in both civil (737 ILCS 5/2 1001) and criminal cases (735 ILCS 5/114-5), although doing so requires additional filings.

The extent to which self-recusal is a problem varies depending on the purported source of bias. For a fairly standard set of potential conflicts (familial relationships, financial interests, pre-judicial legal work on the case, etc.), judges are subject to bright-line, per se recusal rules. In making determinations on these issues, the legal sufficiency of the allegations of bias are generally much easier to produce and evaluate, so less judicial discretion is involved.[107] The judge does technically have the option to refuse removal but doing so would almost certainly result in appeal and an ethical sanction.[108]

For the majority of recusal decisions, however, judges are granted much more decisional freedom. Because it is impractical to predict and outline all of circumstances that might warrant recusal, regimes generally employ broad "catchall" rules that require removal in circumstances in which the judge is sufficiently biased or, in most cases, may be perceived to be biased. In addition to being harder to prove,[109] the standards used to make these self-recusal determinations vary across (and within)[110] jurisdictions and are largely underdeveloped, giving judges such wide latitude that the decision comes down as much to personal choice and perspective as it does to legal standards.[111] This is particularly true in the jurisdictions that use

---

[107] See Melinda A. Marbes, *Reforming Recusal Rules: Reassessing the Presumption of Judicial Impartiality in Light of the Realities of Judging and Changing the Substance of Disqualification Standards to Eliminate Cognitive Errors*, 7 ST. MARY'S J. L. MALPRACTICE & ETHICS 238 (2017) (which explores the distinction between *per se* recusal rules and rules that rely on more judicial discretion. Marbes argues that a separate legal regime should be applied to the two types of recusal decisions.)

[108] Sande L. Buhai, *Federal Judicial Disqualification: A Behavioral and Quantitative Analysis* 69 (2011) (for a review of the cases in the U.S. Circuit Courts in which a decision not to recuse was appealed).

[109] See RICHARD E. FLAMM, JUDICIAL DISQUALIFICATION: RECUSAL AND DISQUALIFICATION OF JUDGES 229 (3rd ed. 2017).

[110] *See, e.g.*, U.S. v. Mikalajunas, 1992 U.S. App. LEXIS 21054 (4th Cir. 1992) (holding that the proper standard for recusal is whether a reasonable person "would" doubt the judge's impartiality), as compared to U.S. v. Cherry, 330 F.3d 658 (4th Cir. 2003) (holding that the proper standard is whether a reasonable pers "might" doubt the judge's impartiality).

[111] See RICHARD E. FLAMM, JUDICIAL DISQUALIFICATION: RECUSAL AND DISQUALIFICATION OF JUDGES 227 (3rd ed. 2017) ("The fact is…that the term 'appearance of bias' has, for the most part, not been defined and, because that is so, courts have not always used that term to mean precisely the same thing.").

One of the likely reasons why this area of law is underdeveloped is that judges often do not need to state the reasons for their recusal on the record when they voluntarily recuse. *See* Patrick A. Woods, *Reversal by Recusal: Comer v. Murphy Oil U.S.A., Inc. and the Need for Mandatory Judicial Recusal Statements*, 13 U.N.H. L. REV. 177 (2015). Similarly, voluntary recusal in a previous case does not necessarily

the appearance-of-bias test, as this standard generally requires the judge to "evaluate the evidence of possible bias in the same manner a reasonable and informed other would."[112]

As some have previously highlighted, the difference in judicial decisionmaking created by these two substantive standards—per se and catchall rules—is substantial and likely necessitates divergent procedural approaches.[113] The current regime of judicial recusal continues to put a premium on providing judges with broad flexibility to determine which cases warrant removal due in large part to the fact that judges vociferously defend their own decisionmaking territory. There is a fear that a full-scale reversion back to the previous recusal regime—an approach hallmarked by automatic recusals and the inherent assumption that judges are biased in many cases—will reflect poorly on the judges and the judiciary more broadly.[114] It is no surprise, then, that the decisions of judges in regard to recusal decisions in specific cases should be—as with their approach to recusal rules and regulation more broadly—partly self-preserving.

ii. The Psychology of Judicial Decisionmaking: Likely the most well-understood—and most discussed—factors that work to disincentivize judicial recusal are the cognitive and psychological difficulties in identifying and overcoming bias and partiality.[115] As we outlined

---

demonstrate the need for recusal in subsequent cases. See Mackey v. U.S., 221 F. App'x 907, 910 (11th Cir. 2007) ("It is simply not true that once recused, always recused..."); Diversified Numismatics, Inc. v. City of Orlando, 949 F.2d 382, 385 (11th Cir. 1991) ("Prior recusals, without more, do not objectively demonstrate an appearance of partiality.");

[112] Melinda A. Marbes, *Reforming Recusal Rules: Reassessing the Presumption of Judicial Impartiality in Light of the Realities of Judging and Changing the Substance of Disqualification Standards to Eliminate Cognitive Errors*, 7 ST. MARY'S J. L. MALPRACTICE & ETHICS 238, 288 (2017).

[113] *See* Melinda A. Marbes, *Reforming Recusal Rules: Reassessing the Presumption of Judicial Impartiality in Light of the Realities of Judging and Changing the Substance of Disqualification Standards to Eliminate Cognitive Errors*, 7 ST. MARY'S J. ON LEGAL MALPRACTICE & ETHICS 238 (2016-2017) (arguing that a two-pronged approach to recusal is necessary to balance the costs of unnecessary disqualification disputes in recusal determinations that involve little judicial discretion against the prominent role that cognitive errors play in recusal determinations that are more discretionary).

[114] *See* CHARLES E. [SIC] GEYH & W. WILLIAM HODES, REPORTERS' NOTES TO THE MODEL CODE OF JUDICIAL CONDUCT 17-18 (2007) (chronicling the heated debate with the ABA regarding the appearance-of-bias standard). *See also* Charles Geyh, Myles Lynk, Robert S. Peck & Toni Clarke, *The State of Recusal Reform,* 18 N.Y.U. J. LEGIS. & PUB. POL'Y 515 (2015) (a panel discussion between a ...on the evolution of recusal standards, particularly in regard to the ABA standards).

[115] For a more substantial review of the importance of judicial psychology as it relates to recusal, see Melinda A. Marbes, *Refocusing Recusals: How the Bias Blind Spot Affects Disqualification Disputes and Should*

earlier, the current judicial recusal regimes in the United States consist of a combination of *per se* recusal rules, which require removal under certain circumstances, and broad catchall provisions, which generally rely almost exclusively on the judge's discretion in determining whether sufficient bias or appearance of bias exists to warrant removal. The discretion provided by this latter category puts the decisionmaking process of the judge front and center, making any cognitive limitations key factors in determining whether and when a judge will actually recuse.

A series of recent studies in psychology have provided evidence that judges are susceptible to many, if not all, of the same cognitive errors and illusions impacting the decisionmaking processes that have long been identified in the wider psychological literature.[116] Judges, for example, are prone to hindsight bias—or the tendency to underestimate the difficulty of past decisions once present outcomes are known[117]—and framing effects—the categorization of difficult decision options as a loss or a gain that depends on how the decision is initially presented[118]—despite their legal training and perception as mechanistic legal arbiters.

Whereas these cognitive errors may arguably lead to the sorts of judicial partiality that might necessitate recusal, the "bias blind spot" is a cognitive error that is particularly relevant to

---

Reshape Recusal Reform, 32 ST. LOUIS U. PUB. L. REV. 235 (2013); Melinda A. Marbes, *Reshaping Recusal Procedures: Eliminating Decisionmaker Bias and Promoting Public Confidence,* 49 VALPARAISO U. L. REV. 807 (2015); Melinda A. Marbes, *Reforming Recusal Rules: Reassessing the Presumption of Judicial Impartiality in Light of the Realities of Judging and Changing the Substance of Disqualification Standards to Eliminate Cognitive Errors*, 7 ST. MARY'S J. ON LEGAL MALPRACTICE & ETHICS 238 (2016-2017); W. Bradley Wendel, *Campaign Contributions and Risk-Avoidance Rules in Judicial Ethics*, 67 DEPAUL L. REV. 255 (2018);

[116] See Chris Guthrie, Jeffrey J. Rachlinski, and Andrew J. Wistrich, *Inside the Judicial Mind*, 86 CORNELL L. REV. 777 (2001) (featuring an empirical study of over 160 federal magistrate judges—one of the first psychological studies on actual judges—testing five common cognitive illusions that had previously been identified to exist among the lay population).

[117] Stephen Hoch and George F. Lowenstein, *Outcome Feedback: Hindsight and Information,* 15 J. EXPERIMENTAL PSYCHOL.: LEARNING MEMORY & COGNITION 605 (1989).

[118] See the famous work of Kahneman and Tversky for a detailed discussion and evaluation of framing: Amos Tversky and Daniel Kahneman, *The Framing of Decisions and the Psychology of Choice*, 221 SCIENCE 453 (1981); Daniel Kahneman and Amon Tversky, *Choices, Values, and Frames,* 30 AM. PSYCHOLOGIST 341 (1984); Amos Tversky and Daniel Kahneman, *Rational Choice and the Framing of Decisions*, 59 J. BUS. S251 (1986).

the ability for judges to make self-recusal determinations because it relates specifically to one's ability to judge one's self. Originally coined by social psychologist Emily Pronin and her associates, the bias blind spot is a descriptor for two related psychological phenomena.[119] First, individuals are systematically apt to underestimate or not even recognize the extent to which they are biased in a particular area. These findings cut to the core questions surrounding judicial recusal, as judges are required to self-determine the existence or perception of bias and will likely not perceive the bias or dismiss it if claims of partiality are raised by outside sources (litigants, their attorneys, or third parties). Second, individuals overestimate the extent to which others are prone to the same bias.[120]

Although it is currently unclear exactly how susceptible judicial decisionmaking and recusal decisions specifically are to this blind spot—no empirical studies have used judges as subjects—leading scholars in recusal law and analysis have appropriately argued that the weight of current evidence suggests that judges are prone to the same errors as the rest of us, so we should presume that they fall prey to the blind spot as well.[121] In the context of recusal, this is particularly problematic because recusal determinations in all the circumstances that do not fall into the enumerated *per se* recusal categories are almost exclusively up to the determination of the judge.

But what of the judges who are successfully able to identify their partiality and sincerely seek to address it, whether through their own inner exploration or through the arguments of one

---

[119] Emily Pronin, Daniel Y. Lin, and Lee Ross, *The Bias Blind Spot: Perceptions of Bias in Self Versus Others*, 28 PERSONALITY & SOC. PSYCHOL. BULL. 369 (2002). See also, Emily Pronin, *Perception and Misperception of Bias in Human Judgment*, 11 TRENDS IN COGNITIVE SCIENCE 37 (2007); Joyce Ehrlinger, Thomas Gilovich, Lee Ross, Peering Into the Bias Blind Spot: People's Assessments of Bias in Themselves and Others, 31 PERSONALITY & SOC. PSYCHOL. BULL. 690 (2005); Irene Scopelliti et. al, *Bias Blind Spot: Structure, Measurement, and Consequences*, 61 MANAGEMENT SCI. 2648 (2015).

[120] This second phenomena may explain the dissonance between what judges say should happen in regard to recusal (see *supra* notes 50 and 51) and what they actually do in practice (see empirical results in *infra* Part V).

[121] Melinda A. Marbes, *Reforming Recusal Rules: Reassessing the Presumption of Judicial Impartiality in Light of the Realities of Judging and Changing the Substance of Disqualification Standards to Eliminate Cognitive Errors*, 7 ST. MARY'S J. ON LEGAL MALPRACTICE & ETHICS 238 (2016-2017)

of the parties? The circumstances surrounding *Caperton* are again relevant in this regard. The West Virginia Supreme Court justice in question was well aware of the donations made by the CEO of Massey Coal and the potential, if not likely, effect that such a relationship could have on his ability to be a neutral decisionmaker in the case.[122] Yet he adamantly declared his ability to ignore the money and stay on the case. Aren't judges particularly well-suited to ignore these sorts of issues or even "de-bias" themselves? Recent findings in psychology are again helpful here. Participants in psychological studies who were informed of the bias blind spot were still prone to believe that they were more able to overcome bias than their peers.[123]

iii. Reputational Costs: Even if judges are successfully able to identify circumstances that lead to partiality or the appearance of partiality, the current ethos of judicial practice frames recusal in such a way that doing so will likely be seen as an abdication of one's role as a judge and may further impugn the judge's colleagues and legal associates.

Recall that in the *Caperton* decision, one of the primary disagreements between the majority and dissenting opinions was the extent to which the justices felt that judges should be presumed to be impartial. Writing in dissent, Justice Roberts argued that the court should be careful to find judges in violation of the "promise" of impartiality, implying that doing so would "impute to judges a lack of firmness, wisdom, or *honor*."[124] In his own brief dissent, Justice Scalia not only expressed a strong presumption of impartiality, but emphasized that overly

---

[122] In the case, Caperton moved for recusal three times, including arguments as to why Justice Benjamin would be biased by the contributions and data from a push poll showing that 67% of West Virginians questioned his ability to be partial (Caperton v. A.T. Massey Coal Co., 129 S. Ct. 2252, 4 (2009)).

[123] Emily Pronin et al., *The Bias Blind Spot: Perceptions of Bias in Self Versus Others*, 28 PERSONALITY & SOC. PSYCHOL. BULL. 369, 375 (2002) (finding that "even the immediate experience of having displayed a particular bias, and then being given an explicitly description of it…, was insufficient to prompt confessions of susceptibility equal to that of one's peers.") (375)

[124] Caperton v. A.T. Massey Coal Co., 129 S. Ct. 2252, 2 (2009) (Roberts, C.J., dissenting) (quoting Bridges v. California, 314 U. S. 252, 273 (1941)) (emphasis added).

inclusive recusal requirements would actually "erod[e] public confidence in the Nation's judicial system . . . ."[125]

The association between recusal and "bad" judging has long been a part of this country's legal culture. William Blackstone (not an American jurist but influential nonetheless) was famously opposed to recusal: "The law will not suppose a possibility of bias or favor in a judge, who is already sworn to administer impartial justice, and whose authority greatly depends upon that presumption and idea."[126] According to Blackstone, a judge cannot recuse himself from a case because, by definition, a judge is someone who would simply never need to recuse. Within this cultural context, to recuse can be seen as an abandonment of the judicial commitment to fairness as opposed to an attempt to maintain it. These cultural motivations are particularly salient when it is one of the participants in the case who raises the issue of bias as opposed to a *sua sponte* motion for recusal from the judge, as "a successful motion to recuse requires the [judge] to admit that he failed in the first instance to adhere to statutory and ethical requirements."[127]

In the context of recusal for campaign donations, the anti-recusal bias is compounded by the fact that almost all judges are trained to be and practiced as lawyers in the communities and political systems where they serve. As Barton describes in his book, *The Lawyer-Judge Bias in the American Legal System*, the shared background and experience between lawyers and judges naturally produces a jurisprudence and judicial culture that favors lawyers.[128] This also means that many of the judges likely participated in the system of campaign contributions from the side of the attorney donor, making it difficult for them to subsequently legitimize the claims that it produces bias by recusing themselves from cases in which an attorney donor is a participant.

---

[125] Caperton v. A.T. Massey Coal Co., 129 S. Ct. 2252, 1 (2009) (Scalia, J., dissenting).

[126] WILLIAM BLACKSTONE, COMMENTARIES ON THE LAWS OF ENGLAND v. 3, at 361.

[127] R. Matthew Pearson, Note, *Duck Duck Recuse?: Foreign Common Law Guidance and Improving Recusal of Supreme Court Justices,* 62 WASH. & LEE L. REV. 1799, 1833–34 (2005).

[128] BENJAMIN H. BARTON, THE LAWYER-JUDGE BIAS IN THE AMERICAN LEGAL SYSTEM (2010).

iv. Institutional Costs: In addition to potential social costs, recusal also brings institutional costs to the judges and the participants in a case. Any judicial transfer will inherently cause at least some administrative burden to the judge and the court system; the case will have to be formally transferred to another judge, requiring some paperwork, a physical transfer of files from one office to another, and some rescheduling. Depending on the jurisdictional procedures and practices, recusal can also trigger a formal review by a peer or presiding judge. Although such costs are likely *de minimis* when recusal occurs at the beginning of the adjudicative process, they can become substantial as the case develops and the filings, deadlines, and scheduled hearings become more closely tied to the specific judge.[129]

Moving from one judge to another can also result in significant loss of invested institutional capital, both to the departing judge and the parties involved in the case. For cases that go beyond the initial stages, judges become familiar with the arguments and participants, meaning that it is difficult for a new judge to catch up mid-case. Similarly, although the legal system relies on and is built to support the general presumption that adjudicative disputes will be resolved in the same way regardless of the individual decisionmaker, it is well known (and supported by empirical evidence) that judges do vary in their procedural approach, legal philosophy, and general disposition,[130] leading to attorneys often tailoring their legal arguments and case-management strategies to the specific judge who is presiding over the case. Having to adjust to a new judge mid-case will likely require additional work hours for the attorneys.

---

[129] See Matthew Menendez and Dorothy Samuels, *Judicial Recusal Reform: Toward Independent Consideration of Disqualification,* Brennan Center for Justice Report (2016), available at https://www.brennancenter.org/publication/judicial-recusal-reform-toward-independent-consideration-disqualification.

[130] See, e.g., Miguel de Figueiredo, Alexandra D. Lahav, and Peter Siegelman, *Against Judicial Accountability,* Working Paper, available at https://ssrn.com/abstract=2989777 (exploring the variations in the time it takes judges to work through cases in the federal court); David S. Abrams, Marianne Bertrand, and Sendhil Mullainathan, *Do Judges Vary In Their Treatment of Race?,* 41 J. LEGAL STUD. 347 (2012) (exploring variations in sentencing behaviors towards defendants of differing racial categories); Sean Farhang, Jonathan P. Kastellec, and Gregory J. Wawro, The Politics of Opinion Assignment and Authorship on the U.S. Court of Appeals: Evidence From Sexual Harassment Cases,*44 J. LEGAL STUD.* 59 (2015).

v. Future Elections: When deciding whether or not they will recuse from a case, elected judges must also consider how their decision might impact future election campaigns. The concern that elected judges might spurn the law in order to increase the likelihood that they are reelected lies at the heart of the debate regarding the legitimacy of popular elections as a method for selecting judges.[131] We have already discussed the empirical research indicating that the legal decisions that the judges make are influenced by the proximity and competitiveness of reelection campaigns,[132] and it is not unreasonable to expect that the "majoritarian difficulty" plays a role in recusal considerations as well.[133]

One might think that the pressures of reelection should actually incentivize higher rates of recusal. The electorate has, after all, shown that they are distrustful of judges raising money for election in large part because they are concerned that the money might impact the judges' decisions in court cases.[134] By actively recusing themselves from cases, particularly those that are more publicly salient, the judges may be able to proactively mitigate some of that mistrust.

---

[131] *See* Amanda Frost & Stefanie A. Lindquist, *Countering the Majoritarian Difficulty,* 96 VA. L. REV. 719, 731 (2010) (arguing that, "elective judiciaries pose a risk to the rule of law, which is compromised whenever a judge's ruling is influenced by majority preferences.").

[132] *See, e.g.*, Joanna M. Shepherd, *Money, Politics, and Impartial Justice*, 58 DUKE L. J. 623 (2009) (finding that judges who are not eligible for reelection behave differently than judges who are eligible); Michael S. Kang & Joanna M. Shepherd, *The Partisan Price of Justice: An Empirical Analysis of Campaign Contributions and Judicial Decisions*, 86 N.Y.U. L. REV. 69 (2011); Sanford C. Gordon & Gregory A. Huber, *The Effect of Electoral Competitiveness on Incumbent* Behavior, 2 Q. J. Pol. Sci. 107, at 107 (2007) (finding that "judges in partisan systems sentence more severely than those in retention systems."); Gregory A. Huber & Sanford C. Gordon, *Accountability and Coercion: Is Justice Blind When It Runs for Office?*, 48 AM. J. POL. SCI. 247 (2004) (finding that judge become more punitive in criminal cases as the next election gets closer); Richard W. Brooks & Steven Raphael, *Life Terms or Death Sentences: The Uneasy Relationship Between Judicial Elections and Capital Punishment*, 92 J. CRIM. L. & CRIMINOLOGY 609, 610 (2002) (finding that the likelihood that criminals convicted of murder are 15% more likely to be given the death sentence during the year leading up to the next judicial election).

[133] Steven P. Croley, *The Majoritarian Difficulty: Elective Judiciaries and the Rule of Law*, 62 U. CHI. L. REV. 689 (1995).

As California Supreme Court Justice Otto Klaus famously put it, "There's no way a judge is going to be able to ignore the political consequences of certain decisions, especially if he or she has to make them near election time. That would be like ignoring a crocodile in your bathtub." (LAWRENCE BAUM, JUDGES AND THEIR AUDIENCES 61 (Princeton University Press ed., 2006)).

[134] *See* James L. Gibson, *Challenges to the Impartiality of State Supreme Courts: Legitimacy Theory and 'New-Style' Judicial Campaigns*, 102 AM. POL. SCI. REV. 59 (2008); James L. Gibson, *'New-Style' Judicial Campaigns and the Legitimacy of State High Courts*, 71 J. POL. 1285 (2009).

It is not clear, however, that this is the direction that electoral incentives will push the judges. A survey experiment conducted by Gibson and Caldeira presented subjects with descriptions of various judicial campaigns, in which some of the judges recused from cases featuring donors and some did not. Although they found that individuals do perceive judges that recuse themselves as more impartial, the overall positive effect was minimal compared to the hit to legitimacy that judges take by accepting contributions in the first place.[135] Additionally, as we discussed above, the act of recusal is often perceived to be an admission of bias, so, as Bam has surmised, "judges might feel that recusing themselves for their campaign statements and conduct would imply that the campaigning itself had been improper."[136]

In addition to being potentially unpopular among the electorate, regularly recusing oneself from cases in which a party or attorney is a campaign contributor is likely going to decrease future donations and, as a result, jeopardize the judge's chances for reelection.[137] Although many, if not most, of such donors contribute to judges simply because they want to help that judge to win election, it would be naive to believe that none of them did so without the hope or the expectation that it would benefit them in the courtroom. An empirical study on the behavior of judicial campaign donors by Miller and Curry lends credence to this intuition. They examined the effect of an Alabama statute that required recusals for donations above a certain dollar amount and found that attorneys and business owners are 81 percent and 40 percent less likely to donate more than the threshold amount, resulting in lower fundraising overall.[138]

---

[135] James L. Gibson & Gregory A. Caldeira, *Judicial Impartiality, Campaign Contributions, and Recusals: Results from a National Survey*, 10 J. EMP. L. STUD. 76 (2013).

[136] Dmitry Bam, *Recusal Failure*, 18 LEG. & PUB. POL'Y 631, at 653 (2015).

[137] For a comprehensive review of the literature linking campaign spending and electoral success, *see* Chris W. Bonneau and Damon M. Cann, *Campaign Spending, Diminishing Marginal Returns, and Campaign Finance Restrictions in Judicial Elections*, 73 J. POL. 1267 (2011) (finding that while the link between spending and winning does exist in judicial elections, it is a nuanced one that depends in part on whether the candidate is an incumbent or challenger).

[138] Banks Miller & Brett Curry, *The Effect of Per Se Recusal Rules on Donor Behavior in Judicial Elections*, 34 Just. Sys. J. 125, at 139-140 (2013)

Correspondingly, Shepherd found that judges who are eligible for reelection are more likely to rule in favor of past campaign contributors than judges who are not eligible for reelection, suggesting that judges rule "in a way that will likely increase the future contributions from interest groups at the time of their next reelection campaign."[139]

B. Disclosure as a Supplement to Recusal

As we highlighted earlier, scholars and policymakers have grown increasingly skeptical of the current system of self-recusal and have presented a number of potential remedies in the form of procedural reform, increased ethical training, and widescale institutional transformation. Of all these suggestions, increased transparency and disclosure of potential conflicts are the most common and often believed to be among the most cost-free and efficacious.[140]

---

[139] Joanna M. Shepherd, *Money, Politics, and Impartial Justice*, 58 DUKE L. J. 623, at 672 (2009). These findings were confirmed in a later study by Kang (Michael S. Kang & Joanna M. Shepherd, *The Partisan Price of Justice: An Empirical Analysis of Campaign Contributions and Judicial Decisions*, 86 N.Y.U. L. REV. 69 (2011)).

[140] See, e.g., Amanda Frost, *Keeping up Appearances: A Process-Oriented Approach to Judicial Recusal*, 53 U. Kan. L. Rev. 531 (2005) ("The proposal discussed here takes this disclosure requirement significantly further by requiring the judge to provide directly to litigants in pending cases any information that might be considered to have an impact on the judge's partiality." (at 583)); James Sample and David E. Pozen, *Making Judicial Recusals More Rigorous*, 46 The Judges' J. (2007) ("One way to increase the odds that litigants will learn pertinent information would be to require judges to disclose, at the outset of the litigation, any facts that might reasonably be construed as bearing on the judges' impartiality.") (at 5)); Deborah Goldberg, James Sample, and David E. Pozen, *The Best Defense: Why Elected Courts Should Lead Recusal Reform*, 46 WASHBURN L. J. 503 (2007) ("In the wake of the White decision, enhanced disclosure might be one of the simplest and most important reforms available." (at 527) Although "[d]isclosure is also an incomplete solution, in the sense that provides only the grounds for disqualification; it does not guarantee that a judge will recuse…" (at 528)); Leslie W. Abramson, *Judicial Disclosure and Disqualification: The Need for More Guidance*, 28 Justice Sys. J. 301 (2007) ("At a minimum, required reports about a judge's financial investment or any other disqualifying circumstance should be readily available both to the support staff in the clerk's office where the judge works and to the litigants and counsel." (at 303) "Two modifications to the new [ABA] Code are essential. First, the Code needs broader disclosure provisions so that litigants believe that the integrity of the judicial system and individual judges is self-evident." (at 308)); David K. Stott, *Zero-Sum Judicial Elections: Balancing Free Speech and Impartiality through Recusal Reform*, 2009 BYU L. REV. 481 (2009) (Specifically discussing recusal for political speech, "Mandatory disclosure would strengthen a currently ineffective method of disclosing past speech." (at 509)); Gabriel D. Serbulea, *Due Process and Judicial Disqualification: The Need for Reform*, 39 PEPPERDINE L. REV. 1109, at 1143 (2011); Thomas M. Susman, *Reciprocity, Denial, and the Appearance of Impropriety: Why Self-Recusal Cannot Remedy the Influence of Campaign Contributions on Judges' Decisions*, 26 J. L. & POL. 359 (2011) ("Whenever a judge is called upon to hear a case—at the trial or appellate level—in which a lawyer or party has been a supporter of that judge and the judge knows (or reasonably should have known) of that support, the judge should issue at the start of the proceeding (1) a statement that sets out the nature and size of the contribution or other financial support known to the judge"

The logic behind these proposals is simple: if judges are not incentivized to pursue self-recusal even when there is a legitimate conflict of interest, we should empower those individuals who *are* properly incentivized—namely the "outside" party and attorney who are likely to be disadvantaged by the conflict. And the primary mechanism for such empowerment is informing interested parties of the conflict via in-court disclosure. Although parties and attorneys are generally not able to unilaterally force the removal of a judge, they can raise the issue through a formal motion for recusal, push for more information from the judge and opposing party, and later appeal the judge's decision if they feel it was incorrect.[141] Even if the outside party decides not to ask a judge to recuse or affirmatively waives a recusal motion submitted by the court *sua sponte*, "[n]othing provides stronger evidence to the parties of impartiality than open disclosure."[142] Without proper disclosure, however, the issue may be unknown to the outside party, and none of (or very few of) the procedural protections will be in place to govern the decisionmaking process that normally exist for adversarial adjudication.[143]

---

and the person making it and (2) an affirmation specifically and in writing that the judge has considered the applicable ethical proscriptions and concluded that the donation will not affect the judge's impartiality in that proceeding. Such a disclosure "statement" will not only alert the parties to the contribution, but also it will focus the judge's attention on the dangers of unconscious reciprocity-based bias." (at 384)); Melinda A. Marbes, *Reshaping Recusal Procedures: Eliminating Decisionmaker Bias and Promoting Public Confidence*, 49 VALPARAISO U. L. REV. 807 (2015) ("[I]n order to make these two reforms regarding who will be the sole or final arbiter in disqualification disputes worthwhile, all jurists and the parties must provide meaningful and timely disclosure of possible grounds for conflict or bias." (at 855) "In addition, the other proposed procedural protections will promote more openness and transparency that will positively affect outcomes in disqualification disputes. The requirement of full disclosure of all interests and connections beyond the more limited disclosures now required by both the jurist and the parties will help ensure that all the relevant information is available to both the litigants and the decision maker." (at 868)).

Even proponents of the current regime feel that increased disclosure has merits beyond simply increasing recusal. See David M. Rothman, CALIFORNIA JUDICIAL CONDUCT HANDBOOK, at 215-16 ("The purpose of the requirement of non-disqualifying disclosure is, in part, to give the attorneys information they would need for an affidavit of prejudice…. The purpose is to reaffirm the integrity and impartiality of the court. Nothing provides stronger evidence to the parties of that impartiality than open disclosure. In addition, because the parties are intimately aware of all the circumstances of their case, they are in a better position to bring to the judge's attention information that might cause the judge to consider recusal.")

[141] The exception to this is in jurisdictions that allow for no-cause peremptory disqualifications, which are discussed further as potential solutions in *infra* Part VI.

[142] Merk & Co. v. Superior Court, 2005 WL 880112 (Cal. Ct. App. 2005), at n. 5 (quoting David M. Rothman, CALIFORNIA JUDICIAL CONDUCT HANDBOOK, at 216.

[143] See Amanda Frost, *Keeping up Appearances: A Process-Oriented Approach to Judicial Recusal*, 53 U. Kan. L. Rev. 531, at 555-56 (2005) (identifying five procedural elements necessary for legitimate

Before continuing in a discussion of disclosure, we note that in the context of judicial conflicts of interest (and particularly those coming from campaign contributions), disclosure is confusingly used to describe a variety of conceptually similar but legally distinct actions. This Chapter is focused on the formal, on-the-record disclosure of potential conflicts of interest to the parties within a legal case (what is sometimes called "in-court" disclosure), but disclosure is most commonly used in reference to the laws and practices surrounding the provision of campaign contribution records to the state election commission by the candidate. Most literature that addresses campaign finance disclosure does so in this regard.[144] Of course, these campaign finance disclosure rules, which exist in some form in all 50 states,[145] are invariably tied to in-court disclosure—it is unlikely that the latter is possible without the former—but it is important to draw a distinction between the two. Additionally, in the recusal context, disclosure can also refer to transparency regarding the reasons that recusal was (or was not) made.[146]

i. In-court Disclosure Rules and Procedures: Unlike recusal rules, which feature prominently in the judicial disqualification statutes and ethical codes, in-court disclosure rules are given little attention in either the laws or the literature. Recall that recusal procedures and rules in the United States generally fall into one of two categories—automatic *per se* recusal statutes and broad catchall provisions. To the extent that states and the federal courts have

adjudication: "(1) litigants, not courts, initiate disputes; (2) the disputes are presented through an adversarial system in which two or more competing parties give their conflicting views; (3) a rationale must be given for decisions; (4) decisions must refer to, and be restricted by, an identifiable body of law; and (5) the decisionmaker must be impartial."

[144] See, e.g., Abby K. Wood, *Campaign Finance Disclosure*, 14 Ann. Rev. L & Soc. Sci. _ (forthcoming 2018).

[145] Stuart Banner, *Disqualifying Elected Judges from Cases Involving Campaign Contributors*, 40 STAN. L. REV. 449, at 471 (1988) ("All fifty states and the District of Columbia require candidates for elective office to file reports disclosing all campaign contributions and, for contributions over a certain amount, the names of contributors.").

[146] See Patrick A. Woods, *Reversal by Recusal: Comer v. Murphy Oil U.S.A., Inc. and the need for Mandatory Judicial Recusal Statements*, 13 U.N.H. L. REV. 177 (2015) (discussing the potential merits of requiring written explanations of why recusals were or were not made).

specific disclosure rules, they also follow that general pattern, although the list of enumerated circumstances requiring disclosure is much shorter.

The most substantial (and well-known) mandatory disclosure requirements exist in the federal courts in relation to financial information. Under Sections 101 and 102 in the Ethics in Government Act, federal officers—including federal judges—are required to report nearly all their sources of income, gifts, property interests, investments, and liabilities on a yearly basis.[147] After a 2006 expose by the Washington Post that described a myriad of ethical conflicts in the federal courts,[148] most courts began to use these reports to create automatic screening procedures so that potential conflicts would always translate into in-court disclosures to the parties (and usually recusal).[149] Even this system, however, relies on full and accurate initial, *ex ante* disclosure by the judge in order for the conflict to be flagged in a given case, and recent study by the Center for Public Integrity found that at least 26 cases from 2011 to 2014 featured a judge who owned stock in a company party in the case.[150]

For other sources of judicial conflict in the federal system and in the fifty percent of states that do include formal disclosure rules and procedures, judges are usually required to inform the parties of relevant, potentially disqualifying information under broad catchall provisions similar to those used for recusal.[151] These provisions are often based on commentary in the ABA Model Code of Judicial Conduct, which suggests a rule that judges disclose any

---

[147] Ethics in Government Act, §§ 101–102, 5 U.S.C. app. 4 (2000).

[148] Joe Stephens, "Ethics Lapses by Federal Judges Persist, Review Finds," WASHINGTON POST (April 18, 2006).

[149] For an example of these plans, see Judicial Council of the District of Columbia Circuit, *Mandatory Conflict Screening Plan*, U.S. Court of Appeals District of Columbia Circuit (accessed August 21, 2018), available at https://www.cadc.uscourts.gov/internet/home.nsf/Content/VL%20-%20Rules%20Policies%20Procedures%20-%20Mandatory%20Conflict%20Screening%20Plan/$FILE/mandatoryconflictplanfinal2.pdf.

[150] Reity O'Brien, Kytja Weir, and Chris Young, *Federal Judges Plead Guilty,* The Center for Public Integrity (accessed August 21, 2018), available at https://www.publicintegrity.org/2014/04/28/14630/federal-judges-plead-guilty.

[151] Leslie W. Abramson (*Judicial Disclosure and Disqualification: The Need for More Guidance*, 28 JUSTICE SYS. J. 301, at 304-5 (2007)) suggests that such codes exist in roughly half of states.

information that the parties or their attorneys "might reasonably consider relevant to a possible motion for disqualification."[152] Some state rules provide an even broader obligation on the judge to disclose, such as in Tennessee courts, where "[a] judge should disclose on the record information that the judge believe the parties or their lawyers *might* reasonably consider relevant to a *possible motion* for disqualification, *even if the judge believes there is no basis for disqualification*."[153]

Because in-court disclosure has primarily been presented as a solution or supplement to recusal as opposed to a topic of study unto itself, it is not well known how these provisions are approached in practice by judges dealing with potential conflicts. On the one hand, the language in the legal provisions require disclosure even when the judge may not believe there is a conflict, as long as the information might be considered relevant to the issue by one of the parties or their attorneys. This would suggest that, at a minimum, disclosure rates should be at least as high as recusal rates and would likely occur much more frequently across a more diverse set of potential conflicts. In practice, however, because the decision to disclose involves even more discretion on the part of the judge than the decision to recuse, many of the same perverse incentives pushing a judge away from recusal likely play into the decision to disclose as well.[154] By definition, disclosure is to make public something that was previously unknown and is possibly problematic, so by introducing a potential conflict via disclosure, the judge is opening herself up to reputational and electoral costs.

---

[152] Model Code of Judicial Conduct Rule 2.11 Comment[5] (2011). This is a change from the previous code, which recommended disclosure that court participants "might consider relevant…" (Model Code of Judicial Conduct Canon 3E(1) Commentary (1999) (Am. Bar Ass'n, amended 2011)).

It is also unclear whether this provision, which states that judges "*should"* disclose, technically requires disclosure or is merely strongly suggestive. See Raymond J. McKoski, *Disqualifying Judges When Their Impartiality Might Reasonably Be Questioned: Moving Beyond a Failed Standard*, 56 AZ. L. REV. 411 (2014), at note 116.

[153] TENN. SUP. CT. R. 10 Canon 3D, Rule 2.11, Comment 5 (emphasis added). See also, e.g., Fla. Stat. Ann. Code of Judicial Conduct Canon 3(E)(1) (2000); Wisconsin Supreme Court Rules 60.04 (4) Comments.

[154] See Leslie W. Abramson, *Judicial Disclosure and Disqualification: The Need for More Guidance*, 28 JUSTICE SYS. J. 301, at 305 (2007).

Disclosures of campaign contributions may be particularly costly in these regards—previous survey work on judicial campaign finance and recusal has shown that public knowledge of donors in the courtroom damages the perception of that judge, an effect that is only partially mitigated by the judge eventually recusing herself.[155] Disclosure may also be understood to be arming the opposition (generally the "outside" party) with more ammunition for a valid recusal motion. This, of course, is one of the primary purposes of disclosure, but as recusal scholar Grant Hammond has noted, "[m]any judges will query why they should hand counsel a stick, with which they can then be beaten."[156] Some jurists have also expressed concern that a disclosure of an innocuous potential bias may damage the litigant's confidence in the judge's ability to be impartial.[157]

C. Why Attorneys Don't Ask Judges to Recuse (And Why Disclosure Doesn't Help)

Just as with the judge's decision to recuse, fully understanding why a party may or may not formally make a motion for recusal, even when the potential partiality has been made known to the outside attorney, requires an exploration of the sometimes perverse incentive structure surrounding the attorney's decision to request recusal.

i. Judges Take Requests for Recusal as Impugning Their Judicial Character: If, as we have suggested above, judges do not like to recuse in part because they see it as a dereliction of their judicial duties, they will certainly not appreciate parties or their attorneys suggesting they do as much.[158] Additionally, because a motion for recusal made by an attorney is necessarily

---

[155] James L. Gibson & Gregory A. Caldeira, *Campaign Support, Conflicts of Interest, and Judicial Impartiality: Can Recusals Rescue the Legitimacy of Courts?,* 74 J. POL. 18 (2012).

[156] GRANT HAMMOND, JUDICIAL RECUSAL: PRINCIPLES, PROCESS AND PROBLEMS, 90 (2009).

[157] See Lord Woolf's comments in *Taylor v. Lawrence,* (2003) QB 528, 549 (Opining that "'judges should be circumspect about declaring the existence of a relationship where there is no real possibility of it being regarded by a fair-minded and informed observer as raising a possibility of bias," at the risk of "unnecessarily undermin[ing the litigant's confidence in the judge.]"

[158] See Penny J. White, *A New Perspective on Judicial Disqualification: An Antidote to eh Effects of the Decisions in* White *and* Citizens United, 46 INDIANA L. REV. 103 (observing, as a judge herself, that "Perhaps human nature causes judges to view disqualification motions as a challenge to their personal integrity. Certainly, no judge, and arguably no person, enjoys being told that he or she is, or appears to be, unfair. It is understandable, therefore, that some (perhaps, many) judges take umbrage at the filing of

done in the absence of the same *sua sponte* motion by the judge, it further suggests that the judge failed in appropriately raising the issue in the first place.[159] Asking for recusal due to campaign contributions—which inherently suggests some form of judicial and electoral impropriety in addition to partiality—is even more likely to offend a judge's sensibilities.

Many, if not most, judges are likely quite modest in their reactions to recusal motions, but given the relative power imbalances between judges and attorneys, the costs of angering some judges can range from mild social discomfort to outcome-changing bias and even professional repercussions for the attorney.[160] The literature and legal cases dealing with judicial recusal are replete with examples showing judges' distaste for recusal motions, including the quote shared at the beginning of this Chapter[161] and a similar sentiment expressed by a judge in the 19th century: "[l]awyers who wanted to try to disqualify a federal judge were 'advised to write out their motion to disqualify on the back of their license to practice law.'"[162] In his dissenting opinion in *Caperton*, Chief Justice Roberts bemoaned the position that recusal motions put judges in, both at an individual and institutional level.[163]

---

disqualification motions." (at 118))

[159] See Richard K. Neumann Jr., *Conflicts of Interest in Bush v. Gore: Did Some Justices Vote Illegally?*, 16 Geo. J. Legal Ethics 375, 392 (2003).

[160] See Melinda A. Marbes, *Reforming Recusal Rules: Reassessing the Presumption of Judicial Impartiality in Light of the Realities of Judging and Changing the Substance of Disqualification Standards to Eliminate Cognitive Errors*, 7 ST. MARY'S J. ON LEGAL MALPRACTICE & ETHICS 238 (2016-2017), at ft. 158) (describing a case in which a judge referred an attorney to the state disciplinary council for filing a recusal motion (see Matthew Mosk & Brian Ross, *'Circus' Continues? Judge Goes After a Lawyer Who Challenged Her Over Controversial Jet Deal*, ABC NEWS (Feb. 11, 2015, 10:45 AM), http://abcnews.go.com/ US/circus-continues-judge-lawyer-challenged-controversial-jet-deal/story?id=28886937) and a case where a judge threatened to report an attorney to the state bar for the same reasons (see *In Re* Cohen, 99 So.3d 926, 940 (Fla. 2012) (per curiam)).

[161] "If you point the barrel of a recusal motion at a Texas judge, make sure it is loaded with a silver bullet." (Jeff Nobles, *Judicial Recusal and Attorney Disqualification: An Ethic for Litigators & Other Aliens in a Strange Land*, Texas Bar College of Legal Education: Advanced Civil Appellate Courses (1999), available at http://www.texasbarcle.com/Materials/Events/1295/66861_01.pdf).

[162] RICHARD E. FLAMM, JUDICIAL DISQUALIFICATION: RECUSAL AND DISQUALIFICATION OF JUDGES at 761 (2d ed. 2007).

[163] Caperton v. A.T. Massey Coal Co., 129 S. Ct. 2252 (2009) (Roberts, J., dissenting) (asking whether "the *judge* get[s] to respond to the allegation that he is probably biased, or is his reputation solely in the hands of the parties to the case." (at 10) And positing that "sometimes the cure is worse than the disease.... I believe that opening the door to recusal claims under the Due Process Clause, for an amorphous 'probability of

ii. Attorneys Are Aware that Judges Will Probably Not Recuse: In addition to accounting for the negative consequences of arguing that a judge is unable to be impartial, the submitting attorney has to adjust the potential benefits of having the judge recuse according to the probability that the judge actually will recuse. Although the study we present in this Chapter is among the first to empirically demonstrate low rates of recusal among judges with potential conflicts of interest, it appears to be well known among practicing attorneys that asking a judge to recuse is not only dangerous but likely fruitless as well. When discussing recusal, one Texas-based attorney stated: "There's been a number of times that I thought it would be best for my client if the judge was recused, but I knew that he wouldn't, so why risk it?"[164]

The little collected research on the jurisprudence of recusal determinations bears these empirical observations out, particularly when it comes to potential conflicts stemming from campaign financing. In addition to *Caperton*, in which a judge declined to recuse despite the CEO of one of the parties having spent more than $3 million on the judge's campaign,[165] judges have made the determination not to recuse in cases featuring campaign donors in Louisiana, North Dakota, Pennsylvania, Nevada, Wisconsin, Georgia, Ohio, Texas, and Florida.[166] Judges have even refused to remove themselves for circumstances as extreme as the donor attorneys hosting the judge's election victory party at their law firm.

iii. Attorneys Are Repeat Players in the Court System: Even in the best-case scenario—where an attorney is aware of a potential conflict, submits a request for recusal, and the judge actually removes herself—the attorney still bears the risk of future prejudice against themselves

---

bias,' will itself bring our judicial system into undeserved disrepute, and diminish the confidence of the American people in the fairness and integrity of their courts." (at 14))

[164] One Texas-based attorney stated that, "There's been a number of times that I thought it would be best for my client if the judge was recused, but I knew that he wouldn't, so why risk it?" (cite)

[165] Caperton v. A.T. Massey Coal Co., 129 S. Ct. 2252 (2009).

[166] See Flamm (2017) *supra* note 119. See also In re Pet'n to Recall Dunleavy, 769 P.2d 1271 (Nev. 1988); Gude v. State, 289 Ga. 46 709 S.E.2d 206 (Ga. 2011); In re Disqal. Of Burnside, 113 Ohio St. 3d 1211, 2006 Ohio 7223, 863 N.E.2d 617 (Ohio 2006) (all cases reviewing a lower-court judge's determination not to recuse due to campaign contributions).

and their clients. In most court systems, particularly smaller state courts, the same attorney-judge pairings are frequent, both because attorneys generally practice within a limited geographic space and because some courts have only a few judges—or just one judge—to whom a case in that jurisdiction can be assigned. In the Wisconsin and Texas data featured in our study, for example, some attorneys argued over 20 cases that were assigned to the same judge over a two-year period.[167]

Attorneys who are aware of a potentially deleterious conflict of interest must then weight the benefits of submitting a motion for a judge to recuse against the likelihood that recusal occurs, the potentially increased bias against the attorney and her party if the judge does not recuse, *and* the potential bias against the attorney and her future parties whether the judge recuses or not. Not only does this incentivize against submitting motions or requests for recusal, it raises significant questions regarding the problem of imputed knowledge and attorney-client conflicts of interest.

iv. Attorneys May Have Donated to Other Judges: When it comes specifically to campaign finance contributions, attorneys may also be disincentivized from pushing for recusal because they themselves have donated to and appeared in front of judges in the past. Although the results of this Chapter's study do not support oft-expressed concerns that every attorney donates to every judge, attorney donations often constitute a significant portion of all the contributions that judges receive, meaning that a non-donor in one case may be a donor in another case. In such a system, strong or repeated motions for recusal could be seen as hypocritical if made by attorneys that are known to have donated to other judges. Successfully persuading a judge to recuse might also set a precedent that would disadvantage that attorney

[167] These were donor-attorneys as opposed to non-donor-attorneys, who's behavior this part is modeling. Because we tracked only cases featuring donor-attorney/judge pairs, we have complete information only on the frequency of those pairings. However, unless the litigation behavior of attorneys who donate are substantially different than attorneys who do not donate, we should expect to see similar rates of appearances in front of a judge over the same time period.

(and her clients) in the future when she argues in front of a judge to whom she donated and from whom she could have received preferential treatment.

v. Judges May Not Disclose: Up to this point, this informal model of attorney behavior has assumed that disclosure has been made and, as a result, the relevant parties were aware of the potential conflict. It is not clear, however, that this assumption is prudent, especially given the nebulous disclosure obligations imposed on judges by the legal regime.[168] As we outlined in *supra* Subpart III.B, the laws and rules dealing with recusal suggest that judges should ostensibly disclose potential conflicts at much higher rates than they recuse due to them, but there is no previous empirical evidence of this, and many of the same factors that disincentivize recusal are relevant in regard to the judge's decision to disclose.[169] Interested parties and attorneys may, of course, avail themselves of contribution data provided to the state election commission by the judge's campaign, but such data may be difficult to obtain (although the empirical study in this Chapter shows that doing so in some states is not prohibitively costly).[170]

IV. The Randomized Experiment

In this Part, we present the results of the ever first field experiment conducted on judges. In the experiment, we identify active Wisconsin and Texas trial-court cases in which one of the listed attorneys had donated to the judge's previous election campaign(s), send a random

[168] See Leslie W. Abramson, *Judicial Disclosure and Disqualification: The Need for More Guidance*, 28 JUSTICE SYS. J. 301, 304-5 (2007) ("In these provisions, a judge should disclose information that the judge believes a party or the party's lawyer might consider relevant to disqualification, a provision that depends on the judge's subjective belief about what a party or lawyer might think is relevant.").

[169] For a valuable analysis of judicial behavior in relation to disclosure more generally, see Scott C. Idleman, *Prudential Theory of Judicial Candor*, 73 TEX. L. REV. 1307 (1995) (which presents a theory of judicial candor (of which disclosure is included), outlining and critiquing nine rationales for candor—accountability, power, quality, authoritativeness, justification, notice, catharsis, progress, and moral duty—and six practical and normative constraints on candor—limited foresight, relative inefficacy, consensus-building, moral exigency, institutional legitimacy, and legal phraseology. While Idleman doesn't address judicial recusal or conflicts of interest stemming from campaign finance, much of his analysis provides helpful framing for a discussion on the judge's decision to disclose).

[170] See the descriptions of our data collection processes in *infra* Part IV.

selection of those judges a letter highlighting the potential conflict and asking for recusal, and then track a number of outcomes related to the decision to recuse.

We begin in Subpart A with an overview of the previous empirical work that has been done on judicial recusal and disclosure. Subpart B describes the institutional structures, election systems, and recusal procedures in the Wisconsin Circuit Courts and the Harris County District Courts. Subpart C outlines the design of the experiment, including the data collection efforts and the implementation of the randomized recusal requests. Subpart D discusses the ethical considerations necessary in an experiment such as this.

A. Previous Empirical Work on Recusal and Disclosure

As we discussed above, the literature expressing skepticism regarding recusal rules in U.S. courts is substantial—scholars, policy makers, and even judges have identified a litany of problems with self-recusal and have called for various levels of reform, ranging from minor procedural adjustments to wholesale institutional restructuring. There has been, however, almost no empirical evidence presented on whether these predictions play out.[171] This empirical lacuna is particularly prominent in regard to conflicts resulting from campaign finance.[172] In fact, in our review of the literature, we found only four studies that reported on recusal rates at all, and only two of them featured an empirical evaluation on judicial recusal as part of the primary analysis.

In one of these studies, Buhai identifies all of the U.S. federal courts of appeals cases from 1980 to 2007 in which one of the parties appealed the District Court judge's decision not to recuse.[173] She provides comprehensive descriptive data on the grounds for which the recusal

[171] Upon reflecting on the array of proposals in this area, Le Epstein questioned "the wisdom of adopting any approach to recusal…in the absence of empirical evidence." (Lee Epstein, *Shedding (Empirical) Light on Judicial Selection*, 74 Missouri L. Rev. 563, at 564-65 (2009)).

[172] While not quantitative, these articles do report on a lot of cases.

[173] Sande L. Buhai, *Federal Judicial Disqualification: A Behavioral and Quantitative Analysis*, 90 OR. L. REV. 69 (2011).

request was made (the trial judge had a personal or professional relationship with an attorney, had a racial bias against the party, had prior rulings on the legal issue, etc.) and the rates at which the lower-court decision not to recuse were reversed (19.2 percent to 0.8 percent, depending on the grounds for recusal).[174] The second study addresses a very different question. In their 2013 article, Miller and Curry take advantage of a change in Alabama law that made recusal mandatory if an attorney or party donated above $4,000 to a judge's campaign.[175] Although the law was never enforced,[176] they found that recusal was 450 percent more likely if the donation threshold had been reached.[177]

The other two studies were focused primarily on whether campaign donations affected case outcomes and only reported on recusal rates as supplementary data. In an empirical evaluation of the Louisiana Supreme Court, Palmer found that in 425 cases judges never recused or even disclosed the donations, some of which exceeded $35,000.[178] A similar study conducted on the Ohio Supreme Court by *New York Times* journalists Liptak and Roberts found that in 215 cases that featured donor parties, judges only recused 9 times, and that "[r]ecusals in cases involving [attorney] contributors were all but unheard of."[179]

---

[174] Sande L. Buhai, *Federal Judicial Disqualification: A Behavioral and Quantitative Analysis,* 90 OR. L. REV. 69, at 70-71 (2011).

[175] Banks Miller & Brett Curry, *The Effect of Per Se Recusal Rules on Donor Behavior in Judicial Elections*, 34 Just. Sys. J. 125 (2013) (their primary findings indicate that the *per se* recusal requirements have a substantial effect on the amount that donors are willing to give to judges).

[176] In 1995, the Alabama legislature passed a rule that required recusal when a party or attorney had donated more than $4,000 to the judge's previous election campaign. However, the law was never cleared by the Department of Justice and subsequently never enforced. See Little v. Strange, 796 F. Supp. 2d 1314 (2011) (in which the U.S. District Court for the Middle District of Alabama dismissed a suit challenging the statute because of a lack of ripeness).

[177] Banks Miller & Brett Curry, *The Effect of Per Se Recusal Rules on Donor Behavior in Judicial Elections*, 34 Just. Sys. J. 125, at 129-30 (2013)

[178] Vernon Valentine Palmer, *The Recusal of American Judges in the Post-Caperton Era: An Empirical Assessment of the Risk of Actual Bias in Decisions Involving Campaign Contributors*, 10 GLOBAL JURIST (2010); Vernon Valentine Palmer & John Levendis, *The Louisiana Supreme Court in Question: An Empirical and Statistical Study of the Effects of Campaign Money on the Judicial Function*, 82 TULANE L. REV. 1291 (2008).

[179] Adam Liptak & Janet Roberts, *Campaign Cash Mirrors a High Court's Rulings*, N.Y. Times, Oct. 1, 2006.

Unlike work on recusal itself, the common proposition that disclosure can augment the current recusal rules is based on a substantial amount of empirical and theoretical research exploring the effects of transparency. None of this research, however, has been directly applied to the realm of judicial disclosure, and as we argue above, the institutional framework in which recusal decisions are made—specifically the relationship between the attorney and the judge—is such that disclosure is likely to be less efficacious.

B. Experimental Setting

i. Judicial Systems: The Wisconsin Court System has a similar structure to most state court systems. Wisconsin's municipal courts oversee common, low-level cases such as traffic violations, ordinance matters, and juvenile crimes. Circuit courts, the state's single-level trial courts, take on the majority of the remaining first-instance cases. Appeals are either taken by the Court of Appeals or—more rarely—by the Supreme Court, Wisconsin's highest court.[180]

The circuit courts, those that are the focus of this study, are tasked with conducting most cases that require judge-supervised trials, including most felony crimes and civil cases. Each county in Wisconsin has a circuit court with a jurisdiction that matches the county's geographical boundary, with the exception of three pairs of neighboring counties with small populations (Buffalo/Pepin, Florence/Forest, and Shawano/Menominee), which share a single circuit court judge. The number of judges in each circuit court varies from 1 (26 counties have just one judge) to 47 (Milwaukee County). In total, there are 249 Wisconsin circuit court judges, divided into 69 circuits. Day-to-day procedure in each circuit court is a mixture of state-mandated rules and individualized county practices. Every circuit court is part of one of 10 judicial administrative districts presided over by a chief judge who is appointed by Wisconsin's

[180]Wisconsin Court System, Court System Overview (2014), available at https://www.wicourts.gov/courts/overview/overview.htm.

supreme court. These chief judges take care of the administrative duties for each district and meet monthly with the other chief judges.[181]

Texas' court structure essentially matches that of Wisconsin except the trial-level courts that handle civil matters with at least $200 in dispute are District Courts as opposed to Circuit Courts. The geographical boundaries of these District Courts are drawn at the county level, with a number of counties sharing one district and the largest county—Harris County—having 60. Each county is part of one of 9 administrative judicial regions.[182] The Harris County criminal and civil courts are housed in the same building, but at any given time, a given district judge is assigned to either the civil docket, the criminal docket, or one of the juvenile or family dockets.

ii. Elections and Campaign Finance: Circuit court judges in Wisconsin are elected to staggered six-year terms. The state holds judicial elections in the spring of every year, including a primary if more than two people run for the nonpartisan position. However, most circuit court elections are uncontested and very few require a primary. In the six- year cycle of elections held from 2009 to 2014, just 22 percent of circuit court elections were contested and only 10 percent of elections included a primary. In most cases, incumbents ran unopposed.

Because Wisconsin circuits are coterminous with counties (with the exception of the three sparsely populated circuits mentioned earlier), they vary widely by population. In the population centers in and around Milwaukee County and in Dane County (which contains the city of Madison), election vote totals range from 50,000 to 100,000 and spending is relatively high. In the smallest counties, by contrast, turnout in judicial elections rarely exceeds a few thousand votes.

Wisconsin law exempts state candidates from filing campaign finance reports if their campaigns raise and spend less than $1,000 in a calendar year. As a result, many judges do not

[181] Wisconsin Court System, Circuit Courts (2014), available at https://www.wicourts.gov/courts/circuit/index.htm.

[182]Texas Judicial Branch, State District Courts – Map (2017) available at http://www.txcourts.gov/media/914401/district-court-map-sept-2017.pdf.

file campaign contributions reports with the Wisconsin Government Accountability Board. By necessity, our campaign finance data consists of the 113 winning judges in the 2009-2014 cycle that did file reports with the state, many of whom ran in contested elections. Table 1 provides descriptive statistics from the campaign finance data submitted by these winning judicial candidates.

District Court judges in Texas run in elections every four years, which are staggered depending on when the judge's district was organized. Any judge who intends to receive more than $500 in contributions is required to report campaign finance data,[183] and Texas judges have consistently reported among the highest dollar amounts in the country.[184]

In an attempt to quell the rising cost and potential for conflict that judicial campaigning presented, Texas passed the Judicial Campaign Fairness Act,[185] which places restrictions on the amount of money organizations and individuals can contribute to judicial campaigns. Although candidates for statewide judicial office (namely the state supreme court) can accept individual donations of up to $30,000, the amount that District Court judges (the focus of our study) can raise depends on the size of their judicial district. Those with districts of over 1 million, between 250,000 and 1 million, and less than 250,000 are restricted to donations of less than $30,000, $15,000, and $6,000, respectively.[186]

iii. Recusal and Disclosure Procedure: Recusal procedure in Wisconsin Circuit courts is regulated primarily by Supreme Court Rule (SCR) 60 and Section 757.19 of the Wisconsin Statutes. Although Section 757.19 is technically the "mandatory" disqualification statute,[187] both

[183] Texas Ethics Commission, Campaign Finance Guide for Judicial Candidates and Officeholders, available at https://www.ethics.state.tx.us/guides/JCOH_guide.htm#RequirReports.

[184] Daniel Becker & Malia Reddick, *Judicial Selection Reform: Examples from Six States*, American Judicature Society Report, available at http://www.judicialselection.com/uploads/Documents/jsreform_1185395742450.pdf.

[185] Texas Election Code Section 253.165(b).

[186] Texas Ethics Commission, Campaign Finance Guide for Judicial Candidates and Officeholders, available at https://www.ethics.state.tx.us/guides/JCOH_guide.htm#RequirReports.

[187] *See supra* note 5 for a discussion of the distinction between disqualification and recusal.

Section 757.19 and SCR 60 govern judicial recusal in Wisconsin courts. Wisconsin does have a Code of Judicial Ethics, but a violation of that code cannot be used as legal grounds for recusal.[188]

After emphasizing the expectation and importance of judicial integrity more broadly, SCR 60.04(4) outlines the circumstances in which a judge is required to recuse, a list that more or less reflects the general approach to these enumerated situations described earlier in this Chapter: personal bias or knowledge of the case, previous legal work on the case (including work as a judge in another court), relation to participants, financial interest in the case or the subject matter, and any public statements made that commit the judge to a decision at issue in the case.[189] Section 757.19 has a very similar set of recusal requirements, except that it also mandates recusal when a judge is a party or material witness in the case.[190]

Both SCR 60.04(4) and Section 757.19 also includes broad catchall provisions for when recusal is necessary even if the specific, enumerated rules do not apply. SCR 60.04(4) declares that a judge shall recuse herself when, "reasonable, well-informed persons knowledgeable about judicial ethics standards and the justice system and aware of the facts and circumstances the judge knows or reasonably should know would reasonably question the judge's ability to be

---

[188] State v. Carviou, 154 Wis. 2d 641, 454 N.W.2d 562 (1990).

[189] SCR 60.04(4)(a-f) reads in whole: "**(a)** The judge has a personal bias or prejudice concerning a party or a party's lawyer or personal knowledge of disputed evidentiary facts concerning the proceeding. **(b)** The judge of an appellate court previously handled the action or proceeding as judge of another court. **(c)** The judge served as a lawyer in the matter in controversy, or a lawyer with whom the judge previously practiced law served during such association as a lawyer concerning the matter, or the judge has been a material witness concerning the matter. **(d)** The judge knows that he or she, individually or as a fiduciary, or the judge's spouse or minor child wherever residing, or any other member of the judge's family residing in the judge's household has an economic interest in the subject matter in controversy or in a party to the proceeding or has any other more than de minimis interest that could be substantially affected by the proceeding. **(e)** The judge or the judge's spouse, or a person within the third degree of kinship to either of them, or the spouse of such a person meets one of the following criteria: **1.** Is a party to the proceeding or an officer, director or trustee of a party. **2.** Is acting as a lawyer in the proceeding. **3.** Is known by the judge to have a more than de minimis interest that could be substantially affected by the proceeding. **4.** Is to the judge's knowledge likely to be a material witness in the proceeding. **(f)** The judge, while a judge or a candidate for judicial office, has made a public statement that commits, or appears to commit, the judge with respect to any of the following: **1.** An issue in the proceeding. **2.** The controversy in the proceeding." (comments excluded)

[190] Wisconsin Statutes & Annotations Chapter 757, Section 19.

impartial…."[191] The official comments included in the SCR clarify that even when circumstances do not fit the clear grounds for recusal outlined above, a judge should apply for recusal any time the "judge knows or reasonably should know to raise reasonable question of the judge's ability to act impartially, regardless of whether any of the specific rules in SCR 60.04 (4) applies." Section 757.19 similarly requires a judge to remove herself if she "determines that, for any reason, he or she cannot, or it appears he or she cannot, act in an impartial manner.[192]" As with most of these catchall provisions, both of these rules rely heavily on the objective determination of the judge.[193]

Of particular importance to our study is the fact that the SCR specifically addresses circumstances in which an individual involved in the case had previously contributed to the presiding judge's election campaign. SCR 60.04(7) states that "[a] judge shall not be required to recuse himself or herself in a proceeding based solely on any endorsement or the judge's campaign committee's receipt of a lawful campaign contribution, including a campaign contribution from an individual or entity involved in the proceeding.[194]" The comments on this portion of the code outline the reasoning for this approach, explaining that involuntary recusal due to campaign contributions would discourage democratic participation and "create the impression that receipt of a contribution automatically impairs the judge's integrity.[195]" Although this clearly establishes the fact that judges in Wisconsin are not legally required to recuse for campaign donations, and likely reflects a professional culture in which recusals for this reason are not expected, it is clear from both SCR 60 and Section 757.19 that a judge can still recuse due conflicts stemming from campaign contributions if she feels it is

[191] Wisconsin Supreme Court Rules 60.04(4)

[192] Wisconsin Statutes & Annotations Chapter 757.19(g).

[193] See State v. Pinno, 2014 WI 74, 93 (Wis. 2014) ("The relevant recusal standard in the Wisconsin Statutes is a subjective one, [and is] drafted so as to place the determination of partiality solely upon the judge…").

[194] Wisconsin Supreme Court Rules 60.04(7).

[195] Wisconsin Supreme Court Rules 60.04(7) Comments.

appropriate. It is only within this discretionary arena that the experiment in this Chapter would work.

The circumstances in which disclosure is necessary under SCR 60 exceed circumstances in which the rules require recusal. Judges are expected to "disclose on the record information that the judge believes the parties or their lawyers might consider relevant to the question of recusal, even if the judge believes there is no real basis for recusal."[196] After full disclosure is made, recusal can be waived when all the interested parties formally agree to allow the judge to continue to participate.[197] Any such agreement or discussion is recorded in the case files.

Once a judge has found reason to recuse or one of the parties to the case has asked for recusal, the court must submit an application for judicial assignment to its district's chief judge. The chief judge subsequently reviews the application against the laws listed above, and either approves or disapproves the action. The chief judge may also request clarification or additional information regarding the circumstances that led to the application.[198]

For most of its history, the recusal regime in Texas was based almost exclusively on disqualification conditions in the Article V, Section 11 of the Texas Constitution.[199] The rule in Section 11 is relatively limited, only prohibiting judges from sitting on cases where "the judge may be interested or where either of the parties may be connected with the judge, either by affinity or consanguinity…or when the judge shall have been counsel in the case.[200]"

---

[196] Wisconsin Supreme Court Rules 60.04 (4) Comments.

[197] Wisconsin Supreme Court Rules 60.04 (6): "A judge required to recuse himself or herself under sub. (4) may disclose on the record the basis of the judge's recusal and may ask the parties and their lawyers to consider, out of the presence of the judge, whether to waive recusal. If, following disclosure of any basis for recusal other than personal bias or prejudice concerning a party, the parties and lawyers, without participation by the judge, all agree that the judge should not be required to recuse himself or herself and the judge is then willing to participate, the judge may participate in the proceeding. The agreement shall be incorporated in the record of the proceeding."

[198] Wisconsin Uniform Rules for Trial Court Administration, Section 3: Disqualification of Judges (2007).

[199] *See* John C. Domino, *The Origins and Development of Judicial Recusal in Texas*, 5 Br. J. Am. Leg. Studies 149 (2016) (reviewing the historical development and current status of recusal procedure in Texas).

[200] Texas Constitution Article V, Section 11.

More recently, recusal in Texas has begun to resemble a similar mix of legislative code and ethical rules similar to those that exist in Wisconsin, at least in theory. Rule 18b(1) of the Texas Rules of Civil Procedure enumerate grounds for mandatory disqualification—close relationships with the parties, interest in the case, previous personal legal work on the case, or legal work on the case by a former attorney associate—and rule 18b(2) outlines circumstances in which a judge may voluntarily recuse—a list similar to 18b(1) except that it includes "any proceeding in which…his impartiality might reasonably be questioned…."[201] The Texas Code of Judicial Conduct Rule 3(C)(1) has also been applied to the question of recusal,[202] although there is some question in the jurisprudence as to the particular legal effect that this ethical provision can have on a judge's decision to recuse or not. An appeals court has recently found that a violation of one or more of the recusal provisions in the Code of Judicial Conduct is not sufficient grounds to require recusal.[203]

There are no specific rules addressing recusal for parties or attorneys who have made campaign contributions, meaning that motions to recuse submitted by either the judge or the parties must be based on either the catchall recusal provision in Rule 18b(2) or a more general appeal to the ethical norms of the Texas Code of Conduct. Attempts to require recusal for contributions on strictly legal grounds have been unsuccessful in the past, even in extreme circumstances.[204]

---

[201] Texas Rules of Civil Procedure 18b.

[202] *See, e.g.*, Rhodes v. State, 2011 Tex. App. LEXIS 9431 (2011) (using a reasonable person standard to rule on a recusal motion).

[203] *See* KB Realtron Mgmt. v. Leleon, Number 13-13-00411-CV, 2015 Tex. App. LEXIS 11879 (Tex. App. - Corpus Christi-Edinburg [13th Dist.] Nov. 19, 2015) (finding that a disciplinary action is the proper resolution to a violation of the Code of Judicial Conduct). *See also* Manges v. Martinez, 683 S.W.2d 137 (1984) (where respondent's request for a writ of mandamus to compel voluntary recusal under Rule 18a(d) was denied for lack of legal grounds due to the Code of Judicial Conduct not being sufficient cause for invocation of Rule 18a(d)).

[204] *See* Rocha v. Ahmad, 662 S.W.2d 77 (Tex. App. 1983) (in which the court found that recusal was not required despite two of the three appellate judges having received thousands of dollars in campaign contributions from the appellee's attorneys for previous campaigns. Additionally, the judges' victory celebrations were held at the attorneys' office.)

The procedures for filing and ruling on a recusal motion vary depending on the court in which the case is heard,[205] with the relevant rules for the cases in this experiment coming from Rule 18a of the Texas Rules of Civil Procedure. Under this rule, if the judge finds, on her own volition, that there are grounds for recusal, the court can sign an order of recusal and refer the case to the presiding judge for reassignment. Parties may also submit motions for recusal, although they must do so more than ten days before the next hearing or trial and must include a detailed argument regarding the legal basis for recusal and the evidence available to support the claims in the motion.[206] Orders on motions to recuse are only subject to appellate review if the motion was denied and there was "an abuse of discretion…."[207]

C. Experimental Design

i. Availability of Electronic Case Data: Pursuant to the open records laws of Wisconsin and Texas, all information regarding the affairs of the government—including individual case files—is presumed to be available to the public. Although there are exceptions to this general principle (for example, financially identifying information or the names of vulnerable individuals may be redacted from court records and cases regarding civil commitments or minors may be excluded entirely),[208] the vast majority of cases in both venues are publicly accessible.

The Wisconsin court system provides a number of services that facilitate the gathering of individual case information. Basic, up-to-date information on cases in the circuit courts, such as case type, case status, judge name, and information on the parties and their attorneys can be

[205] *See, e.g.*, Texas Government Code Section 26.011 (governing recusal of the constitutional county courts); Texas Government Code Section 25.00255 (governing the recusal of the probate courts); Texas Government Code 29 (governing the municipal courts); Texas Rules of Civil Procedure Rule 528 (governing the justice courts).

[206] Texas Rules of Civil Procedure 18a.

[207] Texas Rules of Civil Procedure 18a(j)(1).

[208] Wisconsin Circuit Court Access, *Access to the Public Records of the Wisconsin Circuit Courts*, available at http://wcca.wicourts.gov/index.xsl.

accessed online through Wisconsin's Circuit Court Access database (WCCA). We also used Wisconsin's Simple Object Access Protocol (SOAP) to perform court-wide searches for donor-judge pairs. Unlike the WCCA, which is available at no charge, SOAP is a subscription-based data-extraction program that allows for tailored programming and more accurate case identification. As with most courts, the Wisconsin Circuit Courts also provide access to the legal documents submitted in a given case, although these have to be ordered through the court clerks.

We utilize both WCCA and SOAP to identify donor-judge pairs and measure outcomes of interest. We contracted with Court Data Technologies, a privately-owned Wisconsin-based case data service, to write the programming for and run weekly searches of the SOAP database for treatable pairs. Once treatable pairs are identified, we use the individual case information available through WCCA to verify case details, check on case status, and eventually collect outcomes of interest.

Unlike Wisconsin's state-unified court structure, Texas courts function on the county level. Although the Harris County District Clerk's online case search system allowed us to track cases featuring donor-judge pairs,[209] we were unable to locate an access point for the raw Harris County case data similar to Wisconsin's SOAP and therefore had to manually identify treatable cases using LexisNexis' state court docket search engine. Each week, we ran searches for each judge who had received campaign donations using a list of the last names of self-identified attorneys who had donated to that judge's previous political campaigns provided by Texas' publicly available campaign finance disclosures. Cases verified to feature donor-judge pairs using the criteria detailed below were then tracked using the Harris County Clerk online dockets.

[209] Harris County District Clerk, *Search Our Records and Documents*, available at http://www.hcdistrictclerk.com/edocs/public/Search.aspx?ShowFF=1.

ii. Defining Conflicts: In order to find clear potential conflicts of interest, we focus on identifying contributions from attorneys to judges. The states of Wisconsin and Texas encourage campaigns to collect occupation and employer data from contributors, but many contributions lack this information. For both venues, we placed each verified attorney who gave $200 or more to a sitting circuit court judge into our conflict data set.[210] We refer to each unique pairing of a contributing attorney and a receiving judge as a "conflict pair." We then search newly filed Wisconsin and Harris County cases for these pairs every week. These cases, which involve an attorney who gave money to the presiding judge's campaign, are our "conflict cases."

To identify the Wisconsin cases, we obtained a list of attorneys who are members of the Wisconsin Bar Association and matched the attorney list to the list of judicial campaign contributors. Though we used computer matching as an initial guide to likely matches, we verified each attorney match manually. Since our contribution records went back to 2008, we looked up some attorneys online to see if we could cross reference earlier employers listed in the contributor file with what the Bar Association listed as their current employer in professional biographies or LinkedIn profiles. We then uniquely identified each verified Wisconsin attorney with his or her Wisconsin Bar ID number.

Verifying donor-judge pairs in Harris County using the less data-rich search results produced by the LexisNexis engine was considerably more complicated. Using lists of the surnames of attorneys who had donated to a given judge's previous political campaigns, we compiled weekly sets of cases that might possibly feature a donor attorney. We then manually checked the full attorney names in each of those cases to see if any of them approximately

[210] Note that this amount includes only reported, direct contributions, a limitation that limits the generalizability of our study in a political environment that where "dark money" is common. *See* Abby K. Wood, *Campaign Finance Disclosure*, Unpublished manuscript (2018) (discussing the empirical implications of not including "dark money" in studies dealing with campaign finance).

matched[211] the names from the campaign finance lists, and the case numbers for cases with matching attorneys were entered into the Harris County Clerk's online docket system. Because neither the LexisNexis results nor the online dockets included Bar ID numbers (or any other truly unique attorney identifiers) we used a systematic—but admittedly less accurate—protocol that excluded cases for consideration if there were multiple attorneys in the Texas Bar database with the same or very similar names who could not be differentiated using additional available data such as place of employment or zip code.[212]

iii. Defining Treatable Cases: Although our data collection method gives us access to information on all types of cases, we have excluded a number of different case types from our study.[213] Many of these case types are not treated because they are inherently non-contentious or simply matters of paperwork (i.e. legal name changes, adoption proceedings, and paternity acknowledgement). We felt that other cases, such as small claims and certain types of probate cases, were so uncontroversial that judges were less likely to take our letter treatments seriously. Restraining order/injunction cases were excluded because, by their nature, the most important portions of such proceedings are finished before we can treat the judges. As noted above, we omit all criminal cases.

iv. Assignment Procedure: Over the course of the study, we identified over 700 cases (398 from Wisconsin and 307 from Harris County) that featured attorneys who had donated to

[211] Any names that were similar in spelling or had similar components (e.g. middle name matched first name, initials were similar, name looked like it could be a nickname for another name [Rusty = Russel, Tony = Anthony]) were considered "approximate" matches.

[212] For the full protocol, see Appendix: Harris County Attorney Verification Protocol.

[213] The list of excluded case types in Wisconsin are as follows (Wisconsin case codes in parentheses): All criminal cases (34001,34003), all small claims (31001-31010), restraining orders/injunctions (30704, 30708-30711, 30713, 30709-30711), adoptions (40401, 40403), support/maintenance actions (40402), paternity acknowledgement (40503), informal and ancillary proceedings (50102, 50103), summary assignments and settlements (50105, 50106), determinations of descent (50109), wills filed/filed for safekeeping (50110, 50111). A full list of all Wisconsin Circuit Court case types can be found at http://www.wicourts.gov/courts/circuit/circuitcodes.htm.

All similar case types were also excluded in Texas, in addition to expunction of records and garnishments.

the presiding judge's political campaign. Of these cases, 228 were excluded from our experimental sample because they were of "untreatable" case type, the judge had been presiding over the case for too long, or we could not verify that the attorney on file was the same individual as the attorney from the campaign finance records, resulting in 472 treatable cases (270 from Wisconsin and 202 from Harris County), or an average of 4 cases per week.

Although all 472 of these cases were eligible for random assignment, not all of them received a treatment. As a general matter, we only assigned cases if they were presided over by a judge who had not previously had a case assigned to either condition. However, because the treatment period of the study lasted almost two years, we felt that judges with cases originally assigned to control (meaning the judge had not received a letter) would become re-eligible for analysis after the original case had concluded or after a 5-month cooling-off period.[214] Any treatable cases presided over by judges 5 months after that judge was assigned to control or after the case had concluded were assigned according to the normal procedures described above. If, on the other hand, the judge presiding over a new treatable case had presided over a case that was assigned to the treatment group in previous weeks (meaning the judge had received a letter), the case was excluded from the assignment process.[215]

If, in a given week, we identified only one assignable case, the case was randomly assigned to the treatment condition with either a 50% or 66.6% probability of treatment.[216] If we identified more than one treatable case in a given week, those cases were assigned in a block using a form of complete random assignment that ensured that the number of cases assigned to

[214] We were able to verify whether a case had concluded using the online docket systems.

[215] Although these cases were therefore not included in the experimental analysis, they provide valuable supplementary data on the base rates of recusal among the broader population of cases, so we include descriptive statistics for them in our analysis, below.

[216] Near the end of the experiment, the probability of treatment was increased from 50% to 66% in order to bolster the size of the treatment condition. As we describe below, we account for this variation in treatment probability using inverse probability weights.

treatment and control were roughly equal.[217] We account for the differing probabilities of assignment in our analysis.

The exception to this general assignment protocol occurred in the first week of treatment in both venues. When the experiment was launched (October 2014 in Wisconsin and April 2015 in Harris County), we had a backlog of treatable cases that were still "fresh" under our criteria for inclusion. Each judge with more than one case in the backlog group was placed into individual blocks, with one of the judge's cases randomly assigned to treatment and the others to control. Judges in the backlog group with only one case were placed into a single block for random assignment. An example of the blocking for some cases is shown in the Appendix: Sample of Assignment Blocks in Wisconsin.

v. Treatment Procedure: With the exception of the first week of treatment in Harris County, where the 13 treatment letters were sent in three waves over the course of two weeks, treatment letters were sent via UPS Next-Day Delivery to the corresponding judge's chambers (see Figure 1, below, for letter text) within a week after a treatable case featuring a judge-donor pair was identified. If chamber-specific addresses were not available, we sent the letter to the county court address, presuming that all mail will be delivered to the judges or their clerks. The treatment letters themselves came from the Center for Electoral Politics and Democracy at Fordham University.

Upon receiving a letter, four of the judges contacted our organization directly. When this occurred, we responded with a simple description of our organization's goals (judicial integrity) and reemphasized our belief in the importance of unbiased judicial decisionmaking, while

[217] By "complete" assignment, we simply mean that instead of assigning each case individually (often called "simple" random assignment), we assigned the weekly group so that the number of treatment case and control cases were pre-determined. For weeks with an even number of cases, equal numbers of cases were assigned to each condition (with two cases, one case was assigned to each condition; with four cases, two cases were assigned to each condition; etc.). For weeks with an odd number of cases, equal number of cases were assigned to each condition with the remainder case assigned to the control condition (with three cases, two cases were assigned to the control condition and one to the treatment; with five cases, three cases were assigned to the control condition and two to the treatment).

clarifying that we will not contact any parties other than the judge herself. To the maximum extent possible, our responses followed an identical protocol across cases. All conversations proceeded in a way that we believe maintained the ethical and methodological integrity of the study.

Figure 1: Treatment Letter Contents

Dear Judge **XXX** :

It has recently come to our attention that you have been assigned to preside over **CASE NAME (CASE NUMBER#)**. What attracted our attention is that Ms. **PARTY NAME's** attorney, **ATTORNEY NAME**, has been a significant contributor to your election campaign. In particular, Ms. **ATTORNEY NAME** gave $**AMOUNT** to your campaign committee.

We are sure that you agree that the administration of justice must be perceived as impartial to maintain public confidence. We have no connection to this case or any of the parties involved, but as citizens concerned about judicial integrity, we believe your involvement in **CASE NAME** may raise concerns due to Ms. **ATTORNEY NAME**'s ties to your campaign. In light of those potential concerns, we respectfully request that you recuse yourself from this case and ask that it be reassigned to a different judge.

Sincerely,

Judicial Integrity Project, Center for Electoral Politics and Democracy

vi. Measurement and Outcomes of Interest: Cases assigned to both treatment and control were monitored for a minimum of two months[218] before outcomes were collected. Outcomes were measured using the online dockets, which included information on the presiding judge, the attorneys involved with the case, and (purportedly) any case filings or activity. Once all (or most) the cases have concluded and the courts been informed of the study, we will review these applications in more detail. In order to ensure accurate measurements and allow for future

[218] The majority of cases were monitored for much longer.

reproducibility, we saved images of the online case docket at the time of assignment and of measurement.

The primary outcomes of interest in this project was whether and when judges recuse and disclose the potential conflict. While in Harris County, judges can simply recuse themselves from cases, judges in Wisconsin must submit an application for recusal (formally called an "Application for Judicial Assignment") to the chief judge in the district. We have spoken with Wisconsin court officials and have been told that all recusals require such an application and that the submission of and reasons behind such applications should be noted in the online case history. For each case we recorded applications for recusal and recusals granted, although because one never occurred without the other, we report them as one outcome.

We define judicial disclosure as any case document or docket activity in which the judge discloses the financial relationship between herself and the attorney. This can occur in the course of recusal itself, through the submission of the treatment letter into the court record, or via a conversation between the judge and the parties.

Because judges might also seek to address conflicts of interest through more informal and inconspicuous means than an actual recusal, we measure judicial transfers (non-recusal changes in judicial authority) and whether the donor attorney left the case. To measure the possible impact that the letters have on case outcomes, we record the number of days before a case is resolved[219] and how the case is resolved (whether by default,[220] settlement/stipulation, dismissal, or judgment after trial). To evaluate the impact that disclosures have on attorney behavior, we record requests for recusals made by any of the non-donor attorneys, including formal motions for recusal.

[219] We define the length of a case as the number of days it takes from the first day in the week in which we identified the case to the day the case is closed based on the online dockets.

[220] In Wisconsin, default is defined as follows by Wisconsin Statutes 806.02: "A default judgment may be rendered… if no issue of law or fact has been joined and if the time for joining issue has expired. Any defendant appearing in an action shall be entitled to notice of motion for judgment."

vii. Estimation: Because the probability of receiving the treatment in our experiment may vary case by case, depending on the number of available cases presided over by the same judge, the data must be weighted in order to eliminate any confounding relationship between the assignment probability in a given week and unobserved factors that may affect outcomes. Following the procedures laid out in Gerber and Green,[221] we reweight the data using "inverse probability weights" (IPW). IPW allows us to obtain unbiased estimates of the average treatment effect using a weighted difference-in-means. We also obtain consistent estimates of the average treatment effect using regression with controls for baseline covariates.

Formally, the weighting procedure may be described as follows. Let Pi be the probability that case i is assigned to treatment. Let Z_i be a binary variable that indicates whether a given case is assigned to treatment ($Z_i = 1$) or control ($Z_i = 0$). The weight (W_i) for each case is given by the formula:

$$W_i = Z_i / P_i + (1-Z_i)/ (1-P_i).$$

Applying the weights in equation (1), we obtain unbiased estimates of the average effect of assigned treatment (Z_i) on the outcome variable (Y_i) using weighted difference-in-means. In our application, the outcome is a binary variable scored 1 if the judge requests to be recused and 0 otherwise. Equivalently, we can use weighted least squares (WLS) regression to estimate the parameters of the model. The advantage of using WLS regression is that the model may be extended to control for pre-treatment covariates, such as the number of months before a judge's next election.

We calculate our p-values using randomization inference. Since we know the actual probabilities of assignment to each treatment condition and control, we can simulate a large number of randomly generated treatment assignments and obtain a distribution of treatment effect estimates from these assignments under the sharp null assumption of no treatment

[221] ALAN S. GERBER & DONALD P. GREEN, FIELD EXPERIMENTS: DESIGN, ANALYSIS, AND INTERPRETATION, at chapters 3-4 (W. W. Norton & Company 2012).

effect.[222] We can then see where our actual estimate falls in this simulated distribution and calculate a p-value to determine how likely we would be to observe this value if there were no true treatment effect. All of the randomization inference p-values are based on 10,000 simulated random assignments. We will account for the multiple comparisons problem using a Šidák correction.

D. Ethical Considerations

Because of the novelty of this field experiment—to our knowledge this is the first randomized experiment conducted on judges in active cases—we feel it is necessary to include a discussion of the ethical considerations made during the design and implementation of the study. Like many field experiments involving human subjects, our project raises important ethical and legal issues. Indeed, because the experiment took place in the court context, a thorough evaluation of these issues is especially paramount.[223]

As a result, we have taken precautions to minimize—and likely eliminate—the possibility that asking a judge to recuse herself from a case could have a profound effect on any of the participants—a possibility that we argue is remote to begin with. First of all, we focus on contested civil cases[224] and exclude all criminal cases. This decision to set aside criminal trials was made partially because of the high stakes of these matters and partially because of the faster speed with which they proceed through the judicial system. Second, in order to minimize the burden of imposing delays and complications on cases in which the substantial proceedings have already commenced, we intervened only in "fresh" civil cases— cases where it appears that the judge has taken no substantive actions. As we discuss below, we monitor cases filed in

[222] ALAN S. GERBER & DONALD P. GREEN, FIELD EXPERIMENTS: DESIGN, ANALYSIS, AND INTERPRETATION (W. W. Norton & Company 2012).

[223] For in-depth exploration of the legal and ethical concerns of conducting and evaluating field experiments in the courtroom, see Dane Thorley and Jacob Kopas, *Experiments in the Court: The Legal and Ethical Challenges of Running Randomized Field Experiments in the Courtroom*, Working Paper (2018), available at https://ssrn.com/abstract=2994298.

[224] Both courts have a fairly elaborate categorization system, allowing us to eliminate whole categories of civil cases.

Wisconsin Circuit Courts and Harris County District Courts on a weekly basis to identify judges who have recently been assigned to a case in which one of the attorneys had previously donated to that judge's political campaign. Third, we do not treat cases during the three months leading up to the 2015 and 2016 Wisconsin elections if the judges in those cases were running for reelection.

Since the average civil case of the sort we consider before both the Wisconsin Circuit Court and the Harris County District Court is generally resolved within six to nine months and few show any evidence of substantial activity by the presiding judge until at least a month or two after assignment, focusing exclusively on fresh cases significantly reduces the likelihood that our intervention could have any impact, adverse or otherwise, on the litigants. Nonetheless, it is possible that the treatment letters—regardless of whether they lead to recusal—may indirectly influence the mentality with which the judge approaches the case. A judge who receives a treatment letter may, for example, not have realized that one of the attorneys had donated to her political campaign—a scenario that we strongly do not believe is the case, given disclosure rules in both Wisconsin and Texas—and subsequently favor that attorney (and by proxy, her client) in a way that she would not have absent the information highlighted in the letter. Conversely, a judge, after being informed of the donor attorney's contributions, may choose to employ some sort of self-imposed counter-biasing, favoring the attorney who did not financially support her election campaign. Although we took these possibilities seriously, we ultimately determined that they were unlikely, and were no worse that the natural state of such cases, where the presiding judges almost certainly knew about the donation.

The effect of the treatment letters on lawyers who have contributed to judges may be a bit longer lasting because they could lead a judge to recuse herself in all cases involving a particular attorney or attorneys. Three things are worth noting about that possibility. First, as we present below, we found no evidence of increased recusals in the cases assigned to the treatment condition, let alone among the other cases that treated judges presided over. Second,

cases featuring donor-judge pairs are surprisingly rare; our tracking of newly assigned cases since July 1, 2014 shows that only a small number of cases (less than 1 percent) feature lawyers who have given $200 or more to the presiding judge. In short, the vast majority of contentious civil cases in both Wisconsin and Harris County do not involve attorneys who have made a substantial donation, or any donation, to the presiding judge. Third, even if a judge were to refuse to preside over any case where a particular attorney represented the plaintiff or defendant, the remedy would simply be to transfer the case to another judge in the same or an adjacent county, which would have a negligible impact on proceedings given the early stages in which our experimental intervention takes place.

There is also a policy-relevant aspect to our decision to focus on fresh cases, for we suspect that a judge would be more willing to recuse herself from a case in which her investment of time and energy is minimal. It is reasonable to suppose that a judge who has already heard some testimony, reviewed multiple filings, and issued orders would be resistant to a call for recusal from an outside party, if only because of the high costs to the state and litigants of bringing the next judge up to speed. Whether requests for recusals are an effective tactic for watchdog organizations remains unknown, but we suspect that such requests are likely to be most effective as applied to fresh cases.

Beyond these ethical and practical considerations, we also solicited a formal legal opinion from a leading law firm with established practice groups specializing in election law, non-profits, and taxation. Their analysis confirmed that there is no legal impediment to an outside group writing to judges with a request that they recuse themselves in a specific case due to contributions made by an attorney in the case. Such information is available for public comment due to the public nature of court proceedings. The only potential issue comes from any public treatments (i.e., letters to the editors or print ads) in the context of elections, as public treatments could be categorized as campaign expenditures for the benefit of one of the candidates. Even though this particular study does not incorporate such public treatments, we

observe an extended blackout period of the three months leading up to the Wisconsin (April 7, 2015) and Texas (November 8, 2016) elections for any judge who has declared for reelection. Judicial elections in both Texas and Wisconsin are held on a rotating basis, so just under a sixth of the judges (who serve six-year terms) participate in elections in any given year. This restriction poses no more than a minor inconvenience.

V. Analysis, Results, and Limitations

Here we present the descriptive data we collected on recusal and disclosure and analyze the results of the randomized experiment. As our models predicted, judicial recusal rates are quite low, even when judges are sent the treatment letter asking them to remove themselves from the cases. Disclosure rates are even lower—not one judge who received a letter disclosed the donation on the court record. And even though the treatment letter had a substantial impact on disclosure rates, attorneys in those cases were no more likely to ask the judge to recuse, supporting the hypothesis that attorneys are unwilling to move for recusal even when they are aware of a conflict that may disadvantage their clients.

The detailed results are presented in Subpart A. In Subpart B, we highlight some important limitations to keep in mind when interpreting the data, focusing specifically on generalizability, experimental assumptions, and measurement error.

A Note on The Harris County, Texas Sample: In October 2016, the Judicial Integrity Project received a letter from a Harris County Courts senior staff attorney, writing on behalf of the Harris County District judges. The letter informed us that a few of judges who had received the treatment letters had informed the court administration. According to this letter (although our reading of the relevant statutes is different), judges are not allowed to consider *ex-parte*

requests for recusal such as ours, and that we "may continue to send these requests…but the response will be the same."[225]

To prevent any further potential contamination, we discontinued the assignment of cases in Harris County. In all likelihood, at least some of the experimental results produced from the Harris County Sample are biased. Although there was no indication in the letter that the administration or the judges are aware that these letters are part of an academic study, any possibility that the judges have discussed the existence of the letters among themselves brings with it the likelihood of "spillover," or contamination between the treatment and control groups. In a perfectly clean experiment, none of the judges in the control group will be aware of (directly or indirectly) the treatment letters because their behavior in regard to their own conflict cases may be indirectly influenced simply by knowing that other judges are being asked to recuse. It is also possible that judges who were assigned to the treatment group but became aware of treatment letters sent to other judges might treat their own letters differently.

Considering all of this, it is unclear whether the null results we see in the Harris County data (see below) are actually reflective of the true state of the world in each treatment condition or the result of this incidental spillover. As a result, we focus primarily on the analysis on the

[225] The letter did not specify the law or rule that imposes such a restriction, although we suspect it may be Texas Code of Judicial Conduct, Canon 3:B(8):

> A judge shall accord to every person who has a legal interest in a proceeding, or that person's lawyer, the right to be heard according to law. A judge shall not initiate, permit, or consider ex parte communications or other communications made to the judge outside the presence of the parties between the judge and a party, an attorney, a guardian or attorney ad litem, an alternative dispute resolution neutral, or any other court appointee concerning the merits of a pending or impending judicial proceeding. A judge shall require compliance with this subsection by court personnel subject to the judge's direction and control.

We were familiar with this restriction before starting the Harris County arm and understood it not to include our treatment letter, both because we are not "a party, an attorney, a guardian or attorney ad litem, an alternative dispute resolution neutral, or any other court appointee" and, to a lesser extent, because the contents of the letter addressed a non-case relationship between the judge and an attorney—an issue not directly related to the legal merits of the case. Our interpretation is, of course, unimportant if the practical interpretation of this rule by the Harris County Court encompasses our intervention.

Wisconsin results, although we do briefly discuss the procedural outcomes of the Harris County data.

A. Analysis and Results

i. Descriptive Data: Over the course of the two-year experiment in Wisconsin, we identified 270 "treatable" cases—civil cases that feature an attorney who donated $200 or more to the presiding judge, were recently assigned to the donee judge, and are of an appropriate case type. 60 of these cases were selected to be part of the randomized experiment, with 32 cases assigned to the control group (hereafter called control cases) and 28 cases assigned to the treatment group (treatment cases).[226] Although the other 210 cases (non-experimental cases) are not featured in the experimental analysis, they are still valuable as an additional source for measuring baseline rates of recusal and disclosure, so we include them in our descriptive analysis. Of the 60 cases included in the experiment, all but one has concluded.

Table 1: Covariate Balance Across Treatment Groups (Wisconsin)

Pre-treatment Covariate	Treatment Group Weighted Mean	Control Group Weighted Mean	P-value (2-tailed weighted t-test)
Judge Gender (Male)	0.753	0.726	0.812
Attorney Gender (Male)	0.791	0.673	0.310
Case Type (Divorce)	0.225	0.210	0.894
Attorney Side (Plaintiff)	0.631	0.730	0.432
Donation Amount ($)	350.534	333.222	0.746
Donation Proportion*	0.031	0.041	0.675

*Donation proportion is calculated as the ratio of the donor-attorney's contribution in dollars over the total amount of contributions in dollars made to the judge.

Based on t-tests weighted by the probability of treatment, judge gender, donor-attorney gender, donor-attorney side (plaintiff or respondent), case type (divorce case), contribution

[226] See *supra* Subpart IV.D for a detailed description of why cases were or were not considered treatable and of why cases were or were not included in the randomized experiment.

amount, and the ratio of the attorney contribution to the total amount raised by the judge from attorneys are all balanced across the control and treatment groups (see Table 1, above).

Attorney contributions in these cases ranged from $200 to $1,000, and the average donation was $331 in the experimental cases and $379 in the non-experimental cases. Not surprisingly, the proportion of donation amount to total money raised by the judge was sometimes quite low (.01% at the lowest), but some donations constituted a substantial portion of the judge's total fundraising (66% at the highest).

ii. Base Rates of Recusal and Disclosure: The only two existing studies that track recusal decisions in cases featuring campaign donors suggest that recusal rates should be quite low—somewhere between 0[227] and 4 percent.[228] Our results confirm this pattern (see Figure 2, below). Of the 210 non-experimental cases, judicial recusals occurred in only 12 of them (5.7%). Similarly, in the 32 experimental cases assigned to the control group, we observed only 2 recusals. Taken together, these results suggest that under normal circumstances, judges recuse in only 5.9% of cases that feature a donor attorney.

But the raw rates of recusal only tell part of the story, as we should be able to see the reasons for the recusals that did occur by reading the recusal paperwork that the judges provided to the chief judges. Court documents show that in one of the two recusals in the control cases, the judge recused because the donor-attorney was retained. In the other case, however, recusal was due to the judge having "a close relationship with the parties." This suggests that our

---

[227] *See* Vernon Valentine Palmer, *The Recusal of American Judges in the Post-Caperton Era: An Empirical Assessment of the Risk of Actual Bias in Decisions Involving Campaign Contributors*, 10 GLOBAL JURIST (2010); Vernon Valentine Palmer & John Levendis, *The Louisiana Supreme Court in Question: An Empirical and Statistical Study of the Effects of Campaign Money on the Judicial Function*, 82 TULANE L. REV. 1291 (2008).

[228] *See* Adam Liptak & Janet Roberts, *Campaign Cash Mirrors a High Court's Rulings*, N.Y. TIMES, Oct. 1, 2006 (reporting that judges in the Ohio Supreme Court recused in only 9 of the 215 cases that featured donor parties. They also reported that "[r]ecusals in cases involving [attorney] contributors were all but unheard of," but did not provide a precise percentage).

main measurement for recusal may be overinclusive and that the true rates of recusal due to the campaign donation is even lower.

Figure 2: Base Rates of Recusal (Wisconsin)

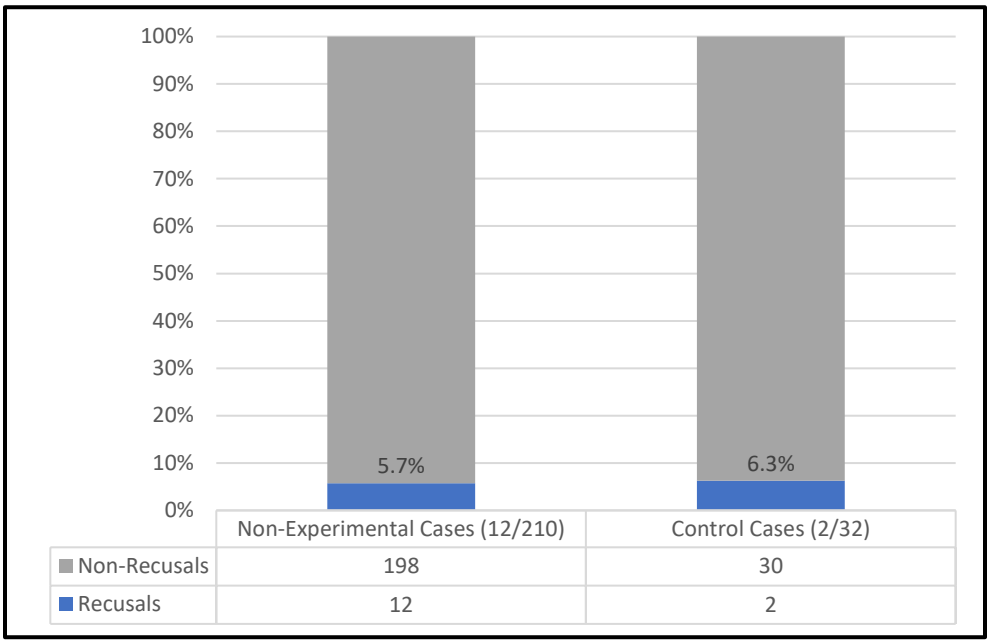
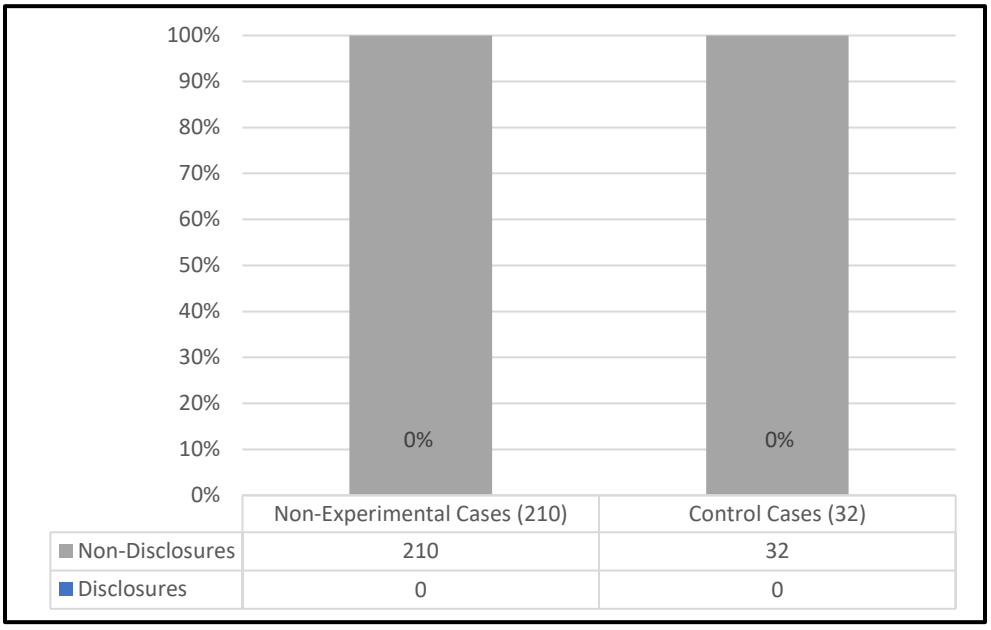


Figure 3: Base-Rates of Disclosure (Wisconsin)



Judges rarely ever recuse in the potential conflict cases that we track in their study, but do they at least disclose the potential source of bias to the parties in the case? Although the presumption in the literature regarding the likelihood of recusal is often that it will be low, disclosure is much more expected than recusal and is arguably required of judges even under a relatively conservative interpretation of the Wisconsin Code of Judicial Conduct. Nonetheless, our study suggests that disclosure is even less common than recusal. In fact, of the 210 non-experimental cases and the 32 control group cases, we did not find a single mention of the contribution(s) made to the judge (see Figure 3, above). This is true even when considering the 14 recusals that occurred between the two groups—while we can assume that some of the recusals that occurred were done at least in part because of the relationship between the donor-attorney and the judge, few of the recusal motions referenced the specific source of bias, meaning that, at least according to our measurement method, disclosure was not made.[229] Interestingly, we did identify some disclosures of potential bias, but they were unrelated to the campaign contributions.[230]

iii. Estimated Treatment Effects on Procedural Outcomes: When designing this experiment, we anticipated that the treatment letters would have a noticeable, moderate effect on the likelihood that judges would recuse from the conflict cases. The results of the experiment suggest otherwise (see Table 2a, below). Judges in the treatment group are only 1.8 percentage points more likely to recuse than judges in the control group, and this small effect is far from statistically significant. Similarly, there was no significant effect of the letter on the likelihood of a judicial transfer or of the donor-attorney's withdrawal.

---

[229] The fact that recusal motions regularly did not express the exact reasons for recusal was initially surprising to use, as we had been told that all recusal motions such include such details. Research by others, however, suggests that this is a common issue. See Patrick A. Woods, *Reversal by Recusal: Comer v. Murphy Oil U.S.A., Inc. and the Need for Mandatory Judicial Recusal Statements*, 13 U.N.H. L. REV. 177 (2015).

[230] For example, one judge "disclosed knowledge of plaintiff's sister," and gave the parties 10 days to determine if they would like the judge to recuse. They subsequently submitted a joint application for judicial recusal.

Table 2a: Estimated Average Treatment Effects on Procedural Outcomes (Wisconsin)

Outcome	Control Group Weighted Mean (n = 32)	Treatment Group Weighted Mean (n = 28)	Estimated ATE (IPW Regression w/ Covariates **)	P-value* (1-tailed)	Standard Errors
Recusals Granted	0.032	0.035	0.018	0.253	0.053
Judicial Transfers	0.161	0.114	-0.007	0.511	0.088
Judicial Disclosure of Donations	0	0.319	**0.337**	**0.000**	0.091
Attorney Withdrawal	0.097	0.117	0.050	0.306	0.079
Any Action* **	0.225	0.390	0.266	0.013	0.115

* The Šidák correction for all seven Wisconsin outcomes combined results in an alpha level of .007. **Coefficients and p-values in bold fall below the Šidák-corrected significance level.**
** Pre-treatment covariates include: judge gender, donor attorney gender, donor attorney side (plaintiff or respondent) case type (binary-divorce case or not), and the combined donation amount given to the judge by the donor-attorney.
***Outcome not included in pre-analysis plan.

There was, however, a pronounced effect of sending judges letters asking for recusal on the propensity for judges to disclose the campaign contribution made by the donor attorney. Although no disclosures were made in the control group, just less than one-third of the judges (32%) in the treatment group either discussed the contribution with the parties in the case, mentioned it in a recusal motion, or included the treatment letter in the court record. This result is statistically significant at the .0003 level (which easily stands up to the increased scrutiny of a Šidák correction).

In the Harris County arm of the experiment, there were also null findings for recusals (0 in the control group, and 0 in the treatment group), transfers (1 in the control group, and 0 in the treatment group), and attorney withdrawal (0 in the control group, and 1 in the treatment

group).[231] Unlike in Wisconsin, however, there was no significant effect on disclosure (0 in the control group, and 1 in the treatment group) (see Table 2b, below).

Table 2b: Estimated Average Treatment Effects on Procedural Outcomes (Harris County)

Outcome	Control Group Weighted Mean (n = 53)	Treatment Group Weighted Mean (n = 17)	Estimated ATE (IPW Regression w/ Covariates **)	P-value* (1-tailed)	Standard Errors
Recusals Granted	.000	.000	.000	1.0	0.000
Judicial Transfers	.000	.000	.000	1.0	0.000
Judicial Disclosure of Donations	.000	.034	.107	.052	.052
Attorney Withdrawal	.000	.059	.024	.043	.030
Any Action* **	.000	.093	.132	.031	.058

* The Šidák correction for all five outcomes tested in the Harris County sample results in an alpha level of .013. **Coefficients and p-values in bold fall below the Šidák-corrected significance level.**
** Pre-treatment covariates include: judge gender, donor attorney gender, donor attorney side (plaintiff or respondent) case type (binary-divorce case or not), and the combined donation amount given to the judge by the donor-attorney.
***Outcome not included in pre-analysis plan.

Because of small sample sizes, we are unable to run robust statistical tests on heterogeneous treatment effects on recusal and disclosure across pre-treatment covariate categories, but we have nonetheless included descriptive statistics on those categorical differences (see Tables 4 and 5 in Appendix). Differences in the treatment effect within subgroups indicate that there is a positive correlation between disclosure and the amount that was donated, the proportion of the judge's total campaign contribution that the donation constitutes, and whether the judge's previous election was contested. Additionally, treatment

[231] Note that the p-values for all of these outcomes are below .05 but do not survive the Šidák correction.

effects on disclosure are larger for male judges and larger in cases where the judge and donor attorney are of the opposite gender, while the treatment effect on both disclosure and recusal is larger in divorce cases relative to other case types. Again, however, a larger experimental sample size is necessary to be confident that these heterogenous treatment effects are more than just chance variation.

iv. Estimated Treatment Effects on Case Outcomes: In addition to inducing recusal, transfers, disclosure, and attorney withdrawal, our letter treatment may also have an effect on the outcomes of the cases. Because of the diversity of civil cases that we included in the experiment, it is difficult to define a "winner" across all case types, but we did measure the number of days it took for the case to conclude[232] and whether the case was concluded through a settlement between the parties.

Table 3: Estimated Average Treatment Effects on Case Outcomes (Wisconsin)

Outcomes	Control Group Weighted Mean (n = 32)	Treatment Group Weighted Mean (n = 28)	Estimated ATE (IPW Regression w/ Covariates **)	P-value* (2-tailed)	Standard Errors
Length of Case (Days)* **	304.853	295.787	6.368	0.463	67.738
Case Settled* **	0.241	0.342	0.030	0.403	0.135

* The Šidák correction for all seven Wisconsin outcomes combined results in an alpha level of .007. **Coefficients and p-values in bold fall below the Šidák-corrected significance level.**

** Pre-treatment covariates include: judge gender, donor attorney gender, donor attorney side (plaintiff or respondent) case type (binary-divorce case or not), and the combined donation amount given to the judge by the donor-attorney.

***Outcome not included in pre-analysis plan.

[232] For the one case that had not concluded, we measured the length of days as if the case had concluded on 10/15/2018.

As we see in Table 3 (above), treatment cases concluded 6.3 days faster than control cases (the raw weighted means, which do not account for the pre-treatment covariates show a larger difference), but this difference is not statistically significant.[233] This difference may be partly driven by the 3 percent increase in the probability that a treatment case is settled by the parties, although that difference is also not significant.

v. Downstream Effect of Disclosure on Non-donor Attorney Behavior: One of the empirical benefits of observing such a large estimated treatment effect on disclosure rates is that the propensity for disclosure effectively becomes a function of the initial random assignment and can be treated as a downstream treatment on subsequent case outcomes. In other words, with some slight adjustments to out estimation strategy we can measure the causal effect of disclosure (specifically the causal effect of a higher probability of disclosure) as if we randomly assigned disclosure itself, something that would be impractical to implement in the original design of the experiment. This then allows us to test the position taken by many in the recusal literature that disclosure can function as a supplement to recusal procedure by informing the non-donor parties of the potential conflict, allowing them to raise the issue of recusal if they deem it appropriate.

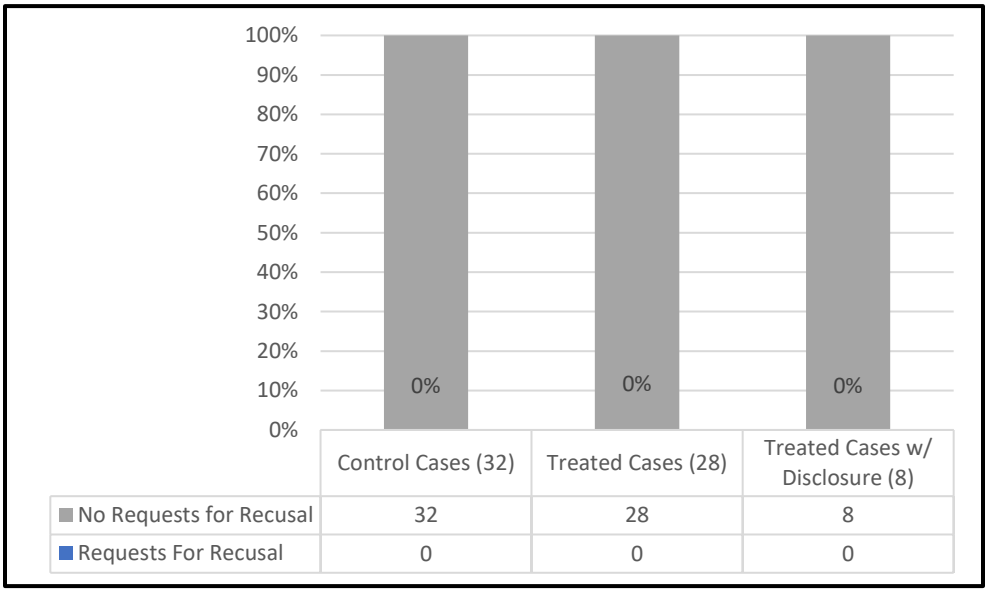
As with the other outcomes of interest, we measured requests for recusal using Wisconsin's online database. We did not find any evidence that attorneys requested recusal in any of the experimental cases, including those in which a recusal was made by the judge. We were told that such requests should be included in the case documents, and based on a review of the non-experimental cases, we have verified that this is true in at least some cases, so not finding any attorney requests is a strong indication that none were made.

With no difference between the control and treatment groups, we do not need to run the more complicated empirical tests for downstream effects to conclude that the increased

[233] For the case outcomes, we used the traditional two-tailed tests because we did not include these outcomes in the pre-analysis plan.

disclosure rates in the treatment group had no impact on the propensity for attorneys to request recusal.[234] Figure 4 (below) presents the rates of requests for recusal in the control and treatment group. We also included the number of requests made just in the cases in which the judge disclosed the contribution to highlight the non-effect that disclosure has.

Figure 4: Rates of Attorney Request for Recusal (Wisconsin)



B. Limitations

The experiment presented in this Chapter provides much needed empirical evidence regarding the frequency and efficacy of judicial recusal and disclosure as well as general judicial and attorney behavior in cases that feature potential conflicts of interest. However, as with all empirical studies, the data presented above must be evaluated with the limitations of our study in mind.

i. Generalizability: Although we believe that Wisconsin and Texas are reasonably representative of the landscape of judicial elections in the U.S., they each feature unique legal and social landscapes that may create different incentive structures for the judges and attorneys

[234] Note that if we were able to conduct downstream calculations, the estimand would be the treatment effect of disclosure on attorney and judge behavior among only those cases in which disclosure was induced.

who work in them. Similarly, we collected data and ran the experiment only on cases in the Wisconsin and Texas trial courts. These cases are of a very different nature than appellate cases, in terms of the role of the courtroom participants and the scope of the conflicts stemming from campaign contributions. Whereas cumulative donations to state supreme court justices can consistently reach the millions of dollars, only a few of the judges in our sample raised more than $100,000, and the attorneys rarely donated more than $1,000. It is not unlikely that the behavioral dynamics influencing both judges and attorneys varies depending on the stage of the judicial process.

The experiment also deals only with conflicts of interest stemming from direct campaign contributions from attorneys. Although much of the consternation regarding the effect that judicial campaign fundraising has on judicial behavior is concerned specifically with attorneys who donate to judges, many of the same threats to judicial legitimacy are at play when it is the parties who donate. *Caperton*, the most prominent U.S. case dealing with recusal for campaign donations features a donor party, not a donor attorney, and would therefore not have been included in our sample. Similarly, because our identification of cases that feature potential conflicts of interest relies exclusively on the publicly reported data that is provided by the judges themselves, this study does not analyze the disclosure and recusal behavior of judges and attorneys involved in cases in which attorneys contributed via indirect spending or undisclosed "dark money," which a recent study by Wood has shown can limit the generalizability of studies dealing with campaign finance.[235]

ii. Experimental Limitations: The study also has a number of potential methodological limitations related to the design and implementation of the experiment. First, the experimental sample is small—32 cases in the control group and 28 cases in the treatment group. Based on the number of cases featuring a conflict in a pre-experimental sample we gathered, we expected

[235] Abby K. Wood, *Campaign Finance Disclosure*, Unpublished manuscript (2018).

a larger number of eligible cases, but the number of cases we could include was ultimately out of our control. As we saw from the results, this did not prevent us from identifying some treatment effects, but it does lower our confidence in the null results observed and prevents a more statistically robust test of the downstream effects of disclosure.

Second, we cannot verify that all of the judges in the treatment group read the recusal requests. It is possible that the letters were intercepted by a judicial assistant and never given to the judges or that the judges simply threw the letters in the trash before reading them. In either case, this would result in a problem called non-compliance in the experimental literature and would produce a biased estimate of the treatment effect. Although we cannot empirically dismiss the possibility of non-compliance, we are not particularly concerned about it in the context of this experiment. Because we tracked the packages sent to the courthouse, we know that the packages were delivered, and because of the 32 percentage-point treatment effect on disclosures, we can surmise that at least that many judges were actually treated.

Third, there may have been experimental spillover (another term of art in the experimental literature) between the control and treatment groups, particularly among the judges in the experiment who are housed in the same courthouse. Spillover occurs when the treatment category of a judge impacts the outcome of interest of another judge. In the context of this experiment, spillover is most likely to occur if a judge in the treatment group discussed the recusal request sent to her with one of her colleagues who was also assigned to either the control or treatment group. As with non-compliance, this sort of interaction between treatment groups could result in biased estimates of the treatment effects, and, as with non-compliance, the extent to which this was a problem in our experiment is difficult (if not impossible) to empirically measure, although as we explained above, we have evidence to suggest that it did occur in the Harris County sample.

iii. Measurement Error: Finally, the result of the study should be evaluated understanding that our measurements of all the experimental outcomes are potentially under-

inclusive of the actual outcomes. This is due in large part to our reliance on online case dockets and formal court documents. We were assured that the online docket system in Wisconsin accurately reflects the totality of events in a given case (excluding any information that is not disclosed for legal reasons), but it is possible that an event that we recorded as a judicial transfer was actually a recusal or that a disclosure of the campaign donation was made in the course of a hearing but not recorded individually on the docket. Although we ordered copies of relevant case events to try and remedy this problem, it is still possible that we missed certain events.

VI. Alternatives

Having analyzed the evidence showing that judges are unlikely to recuse themselves from cases with campaign donors and that simply informing the outside attorneys and parties of that conflict does little to stimulate motions for recusal, we now turn to a discussion of alternative methods for dealing with conflicts of interest. As with much of the previous discussion in this Chapter, our arguments are tailored specifically to conflicts stemming from campaign finance but may be applicable to conflicts similar in nature.

We discuss two proposals. The first, a tightly tailored procedural proposal, comprises a fusion of peremptory judicial challenges and automatic administrative disclosure of attorney donations. We argue that unlike other procedural reforms, this approach accounts for the incentive structure facing both judges and (importantly) attorneys. We also briefly discuss anonymizing judicial campaign contributions as one of the few feasible (albeit still problematic) institutional-level reforms. In generating these alternatives, we adopted much from the ideas and discussions that have been presented in the extant recusal literature, particularly those that do not seek simply to tweak standard self-recusal procedure or remove the system of judicial elections entirely.[236]

[236] In our opinion, a pair of 2007 articles by Goldberg, Sample, and Pozen provide the solutions that are the most aware of the sort of behavioral limitations that we have presented in our discussion. The alternatives

A. Procedural Reform

Our previous analysis suggests that the most prominent procedural approaches for dealing with judicial conflicts of interest appear to fall short for two primary reasons: 1) they afford too much discretion to the judge in question, allowing for the consideration and influence of extra-legal factors that disincentivize self-recusal, and 2) they often rely on motions for recusal that are too costly for attorneys in all but the most extreme circumstances, meaning that transparency and disclosure do little to facilitate the removal of potentially conflicted judges.

The most common procedural or rule-based solutions to the shortcomings of self-recusal either shift the recusal decision to another adjudicator or make recusal mandatory in all or a set of conflict cases. Both of these approaches correctly identify that for recusal to occur at a desirable rate, the recusal decision cannot be made by the adjudicator about whom the conflict is concerned. Nonetheless, neither of these approaches are likely to work when addressing conflicts due to campaign contributions because they either do not account for the incentives of the ultimate decisionmaker or do not account for the incentive structure of the attorneys in the case.

A number of states have sought to avoid the problems associated with self-recusal by requiring third-party recusal determinations. Wisconsin, for example requires the approval of a district's chief judge for any of the trial judges to be recused, although the request for such review has to go through the judge in question.[237] However, independent review of recusal motions is usually conducted by other elected judges who have also received campaign contributions and are therefore subject to many of the same incentives preventing self-recusal. Additionally, if (as we argue in this Chapter) recusal is seen as an abdication of one's duty and

---

discussed in the subsequent literature appear to come mostly from the suggestions. See James Sample and David E. Pozen, *Making Judicial Recusals More Rigorous*, 46 The Judges' J. (2007); Deborah Goldberg, James Sample, and David E. Pozen, *The Best Defense: Why Elected Courts Should Lead Recusal Reform*, 46 WASHBURN L. J. 503 (2007).

[237] Wisconsin Uniform Rules for Trial Court Administration, Section 3: Disqualification of Judges (2007).

judges dislike it when attorneys request recusal, judges are likely not going to require recusal for other judges except in the most egregious circumstances.

If judicial discretion is still an issue even if the recusal determination is exported to outside judges, it stands to reason that we should just remove discretion altogether. Policy makers and scholars have repeatedly argued this point, suggesting that one of the best solutions simply requires including conflicts due to campaign donations into the list of circumstances that trigger automatic, or *per se* recusal.[238] Such a system can make recusal mandatory for cases in which a donation of any amount is made or, as the ABA Model Code has suggested, make recusal automatic for donations that are of a certain amount or proportion of a judge's total fundraising.[239]

Although automatic recusal solves for the incentives of judges and outside attorneys—behavioral incentives play no role in the recusal decision if there is no discretion involved—and ensures that judges will not hear cases in which a participant is a campaign donor, it creates new concerns. Unlike the other conflicts included in *per se* recusal statutes (e.g., familial relationships or previous professional experience with the case) the existence of conflicts stemming from campaign contributions is in the control of the donor, thereby allowing for the strategic creation of the conflict. Instead of donating to judges to support their candidacy or maybe garner favoritism from that judge, attorneys can instead "pay-not-to-play" by donating to judges they perceive as unfavorable or unfriendly.

---

[238] See Gabriel D. Serbulea, *Due Process and Judicial Disqualification: The Need for Reform*, 39 PEPPERDINE L. REV. 1109, at 1143 (2011). See also, Donald L. Burnett, *A Cancer on the Republic: The Assault Upon Impartiality of State Courts and the Challenge to Judicial Selection*, 34 FORDHAM URBAN L. J. 265 (2006) (suggesting *per se* recusals as a solution for non-recusal in the campaign speech context).

[239] Model Code of Judicial Conduct Canon 3E(1)(e) (1999) (Am. Bar Ass'n, amended 2011). Rule 2.11 calls for automatic recusal when "The judge knows or learns by means of a timely motion that a party, a party's lawyer, or the law firm of a party's lawyer has within the previous [insert number] year[s] made aggregate contributions to the judge's campaign in an amount that is greater than $[insert amount] for an individual or $[insert amount] for an entity [is reasonable and appropriate for an individual or an entity]…."

In Utah, for example, recusal is required if any of the parties or attorneys has donated more than $50 to the judge.[240] This allows enterprising attorneys to avoid certain judges for a small one-time fee, behavior that we have been told in informal interviews is, at the very least, contemplated by attorneys working within these states. To combat this, the rules could increase the threshold at which recusal become automatic so that it is high enough to deter such strategic play, but this would inversely decrease the capture of conflicts of interests that the rule was designed to address. Similarly, others have suggested that allowing the opposing side to waive automatic recusal could prevent this strategic behavior,[241] but that would put the burden of identifying gamesmanship on the non-donor party and attorney, who are the worst equipped for making such a determination.

We posit that a more tenable procedural approach to campaign donors in the courtroom is a combination of limited, no-cause peremptory challenges paired with mandatory disclosure by the administrative court system. A good number of states—most prominently California—already allow for parties to submit one no-cause peremptory challenge in each case as long as they do so within the first weeks of the initial case assignment.[242] These challenges result in an automatic, no-questions-asked transfer to another judge that generally occurs so early in the process that the judges are unaware that the case was ever assigned to them in the first place, avoiding much of the potential cost of judicial retribution or bias in the attorney's future cases in front of that judge.

On its own, a proposal calling for peremptory challenges is not wholly novel—legal scholars have previously seen its potential as a solution for the failings of the current recusal

---

[240] Utah Supreme Court Code Rule 2011(A)(4).

[241] See Deborah Goldberg, James Sample, and David E. Pozen, *The Best Defense: Why Elected Courts Should Lead Recusal Reform*, 46 WASHBURN L. J. 503 at 529-30 (2007).

[242] In total, there are 17 states that allow peremptory challenges, although the circumstances in which the challenges are allowed vary. Such states are: Alaska, Arizona, California, Idaho, Illinois, Indiana, Minnesota, Missouri, Montana, New Mexico, Nevada, North Dakota, Oregon, South Dakota, Washington, Wisconsin, Wyoming (RICHARD E. FLAMM, JUDICIAL DISQUALIFICATION: RECUSAL AND DISQUALIFICATION OF JUDGES at 789-822 (2d ed. 2007)).

regime.[243] When considering such a system, some have expressed concerns that without proper restrictions, attorneys may make challenges in the middle of a case if they perceive the outcome to be developing against the interests of their client. This would result in substantial administrative burdens to the court system and the opposing party.[244] In 2010, for example, government attorneys in California used the system to boycott judges who were perceived to be too light on crime by coordinating office-wide policies to skip those judges anytime they were assigned to a case.[245] To allay some of these concerns, this Chapter proposes a peremptory challenge regime that is limited only to cases in which the conflict of interest (or potential conflict) exists—a certain amount of total campaign donations to the presiding judge in the case of this proposal—and can be exercised only by the disadvantaged side, or the side from whom the conflict does not stem.[246]

As the empirical study in this Chapter shows, however, judges are unlikely to disclose campaign donation to the non-donor attorneys unless prompted to by a third-party intervention, so the attorneys may not know to make the challenge even if it is available. Additionally, even if the judges did disclose, the judge would likely realize why any subsequent peremptory challenges were made, re-introducing the behavioral structure that disincentivizes attorneys from asking for recusal. As a result, we also propose that the peremptory challenges be paired with mandatory non-judicial disclosure. The most feasible mechanism for this sort of disclosure would likely be the inclusion of attorney or party contributions in the initial court

---

[243] See Peter A. Galbraith, *Disqualifying Federal District Judges Without Cause*, 50 Wash. L. Rev. 109 (1974); ALAN J. CHASET, DISQUALIFICATION OF FEDERAL JUDGES BY PEREMPTORY CHALLENGE (1981); Deborah Goldberg, James Sample, and David E. Pozen, *The Best Defense: Why Elected Courts Should Lead Recusal Reform*, 46 WASHBURN L. J. 503 (2007).

[244] See Debra Lyn Bassett, *Judicial Disqualification in the Federal Appellate Courts,* 87 IOWA L. REV. 1213, at 1251-56 (2002) (providing an exploration of the common objections to a peremptory challenge regime, specifically applied to the federal appellate court context).

[245] See, e.g., Michelle Quinn, "District Attorney's Boycott of a Judge Raises Issues," NY TIMES (March 20,2010).

[246] If the subsequent judicial assignment also results in a donor conflict, the disadvantaged side can again ask for the judge to be removed.

documents. This system would naturally come with costs to the court administrative system, as it would need a database in order to identify potential conflict pairs.[247] However, such information is, by law, publicly available and creating such a system should be well within the means of court administrations, something that we can attest to after independently identifying conflict cases using the Wisconsin and Texas courts' own systems.

An accurate court record of donors will ultimately rely on accurate and full disclosure of the campaign finance data by the judge, which brings with it some concerns and limitations. First, it is clear that most campaign finance disclosure laws are inadequate at identifying all of the financial relationships that are created by donations, as corporate contributions (including those made by law firms) and dark money do not produce clear connections between individuals and the candidates.[248] Second, the administrative body will still need to rely on the judge, or at least the judge's campaign committee, to be aware of and report all direct contributors. Judges have raised this concern as a reason against adopting both mandatory disclosure and *per se* recusal for campaign donations,[249] arguing that the burden of actively knowing and being able to identify all donors is too onerous. However, because the system we propose puts the onus of disclosure in a given case on the court administration, it would only require an initial report by the judges, which is something they already do as part of the required election law in their state.

B. Institutional Reform

Many scholars and policy makers concerned with the potentially detrimental influence of campaign donations on judicial behavior have focused their efforts on culling conflicts at their genesis by reforming (or removing) judicial elections. The ongoing debates regarding the

[247] See Deborah Goldberg, James Sample, and David E. Pozen, *The Best Defense: Why Elected Courts Should Lead Recusal Reform*, 46 WASHBURN L. J. 503 at 527 (2007); Gabriel D. Serbulea, *Due Process and Judicial Disqualification: The Need for Reform*, 39 Pepperdine L. Rev. 1109, at 1145 (2011).

[248] See Abby K. Wood, *Show Me the Money: 'Dark Money' and the Informational Benefit of Campaign Finance* Disclosure, USC Law Legal Studies Paper No. 17-23 (August 30, 2017); Abby K. Wood, *Campaign Finance Disclosure*, 14 Ann. Rev. L & Soc. Sci. _ (forthcoming 2018).

[249] See, e.g., Charles Geyh, et al., *The State of Recusal Reform*, 18 N.Y.U. J. Legis. & Pub. Pol'y 515 (2015).

respective virtues and vices of judicial elections and judicial appointments is rich and nuanced, and this Chapter does not provide theoretical or normative contributions on that front. We do stress, however, that elections as a method of judicial selection is a consistently popular phenomenon that is the result of institutional dynamics and public attitudes that are still at play in the current political milieu.[250] As a practical matter, then, judicial elections are not going away any time soon and so an attempt to abolish them does not serve as a practically viable approach for addressing judicial conflicts—at least not in the short term. We instead consider more focused, incremental institutional solutions specifically addressing the problems stemming from judicial campaign finance.[251]

One such approach is anonymized donations, where the identity of donors is blocked from the judge. The logic behind anonymization is straightforward: if judges are unable to know the identities of those who donated to their campaigns, they will be unable to favor (knowingly or unknowingly) those individuals in the courtroom. Understanding this to be the case, the public would no longer harbor the suspicion of *quid pro quo* courtroom relationships, and the threat to judicial legitimacy would be minimized. Anonymous donations would ideally render ex-post procedural solutions such as recusal largely unnecessary while still allowing the public—particularly the legal community—to financially support individual judges.

Many states currently have or have experimented with legal reform that is built on the premise of creating distance between judges and their donors. Most states with judicial elections, for example, restrict or ban direct solicitation by judicial candidates, including in-

---

[250] *See* Debra Cassens Weiss, *55% of Potential Voters Support Judicial Elections for State Judges*, ABA JOURNAL (Oct. 21, 2008 5:08 PM), available at http://www.abajournal.com/news/article/55_of_potential_voters_support_judicial_elections; Justice at Stake Campaign, *Frequency Poll of American Voters* (2001), available at http://www.justiceatstake.org/media/cms/JASNationalSurveyResults_6F537F99272D4.pdf (finding that 54% of American favor selecting judges through popular elections over a system of merit selection).

[251] We agree in large part with Geyh, who believes that judicial appointments are most compatible with preserving an independent judiciary but argues that judicial elections are simply too popular and too beneficial to judicial legitimacy to be able to discard them (see Charles Gardner Geyh, *Judicial Selection and the Search for Middle Ground,* 67 DEPAUL L. REV. 333 (2018)).

person requests and even personally signed letters.[252] Judges in these systems are still able to raise money, but the actual exchanges can only occur between donors and designated non-candidate representatives. Similarly, a few states have implemented publicly financed judicial campaigns, where judges are provided with set amounts of campaign funds provided by the state.[253]

In practice, however, judges in these systems are still not fully insulated against the influence of donations. In states that prohibit direct solicitation, judges are generally still aware of who their financial supporters are and are usually able to directly communicate with donors about their contribution after the donation has been made.[254] Furthermore, the states that have or have had public financing are more appropriately understood as public-private hybrids, where participation is either entirely optional—as in New Mexico—or judges running publicly funded campaigns are still allowed to raise some level of private donation. Nonetheless, these reforms are valuable in that they have shown that there is both a public appetite for such restrictions and, importantly, that the U.S. Supreme Court has been more deferential regarding state efforts to regulate judicial elections than similar efforts in the executive or legislative context.[255]

Following Ayres and Bulow,[256] we suggest that for anonymization to work, the administrator accepting the donations should be independent of the judicial candidate or her

---

[252] Brief for the ABA as Amicus Curiae, p. 4, Williams-Yulee v. The Florida Bar, 135 S. Ct. 1656 (2015) (claiming that 30 of the 39 states that elect judges have adopted restrictions on personal solicitation of campaign funds by judges).

[253] See American Bar Association Standing Committee on Judicial Independence, *Public Financing of Judicial Campaigns: Report of the Commission on Public Financing of Judicial Campaigns* (2002) (available at https://www.americanbar.org/content/dam/aba/migrated/judind/pdf/commissionreport4_03.authcheckdam.pdf).

[254] The ability for judges to send personalized thank you letters while being unable to personally solicit funds was an inconsistency discussed at length by the dissent in Williams-Yulee v. The Florida Bar, 135 S. Ct. 1656 (2015).

[255] See *Id.*

[256] Ian Ayres and Jeremy Bulow, *The Donation Booth: Mandating Donor Anonymity to Disrupt the Market for Political Influence*, 50 STANFORD L. REV. 837, at 870-75 (1988). See also Stuart Banner, *Disqualifying Elected Judges from Cases Involving Campaign* Contributors, 40 STANFORD L. REV. 449, at 474-76 (1988)

campaign committee, and the system should be mandatory for all candidates. Those authors posit that the ideal system would utilize blind trusts established by the candidates and operated by private trust companies, but we feel that trusts run by a centralized governmental body would be less complicated and equally sufficient. To help ensure anonymity, donors would be prohibited from discussing the donations with candidates, and the administrator would report only on net contributions, as opposed to individual donation amounts.

As with nearly all large-scale institutional reforms, this sort of system is not without limitations and drawbacks. If judges do not know who donated to their campaign, judicial contributions will probably go down in both frequency and amount. Although such a decrease would be illustrative of the fear that much of the money involved in judicial politics is spent with the express intention of signaling support to and garnering favor from judges, judicial candidates will naturally bemoan smaller war chests. Additionally, it would be naïve to believe that any system of anonymization would act as a total bar to information. Motivated donors—potentially those that are most likely to engage in the sorts of *quid pro quo* relationships that are most concerning—will find ways to inform judges of their contribution[257] or shift to less regulated forms of support. Additionally, any system that would credibly keep the identities of donors from judges would, by necessity, also keep that information from the public, reversing the move towards transparency that has motivated much of the recent campaign finance reform.

VII. Conclusion

This Chapter makes two principal claims regarding the recusal regime predominately used in U.S. courts: first, that judges will not be incentivized to recuse from cases in which they

(reviewing efforts by the ABA and some states at shielding judges from donor information); Dimitry Bam, *Making Appearances Matter: Recusal and the Appearance of Bias*, 2011 B.Y.U L. REV. 943, at 986-88 (2011).

[257] Ayres and Bulow argue that such communication would necessarily be "cheap talk" because without the ability to prove a donation was made, anyone could claim financial support, and the judges would not know whether any purported donor was telling the truth. *See, Id* at 855-56.

may be partial, and second, that attorneys will not be incentivized to ask them to recuse from these cases even if they are aware of the potential source of bias. The experimental evidence presented above is consistent with both claims.

It is clear from the data that judges almost never recuse from cases in which an attorney had donated to their political campaign. The rate of recusal within the 32 cases featured in the control group of the experiment as well as within the 210 non-experimental cases that feature these potential conflicts is just above five percent. Surprisingly, these rates did not increase (at least not in a significant amount) even when the judges were sent a somewhat obscure letter that identified the donation and asked the judge to recuse.[258] Additionally, roughly half of the recusals that we did observe were purportedly done for reasons unrelated to the donation (relationships with the parties, generally).

If recusal procedures and rules are meant to lead judges to recuse in cases of with conflicts, the failure of judges to recuse from the cases in this study is strong evidence that the rules are not fulfilling their purpose. Many who have written on recusal have guessed that the current recusal regime does a poor job of removing judges from cases in which they might be biased, but it is unlikely that any except for the most skeptical would have expected such low levels of removal. If judges themselves are to be believed, we should expect recusal in roughly half the cases: over 40 percent of judges believe that campaign contributions influence their (or at least their colleagues') decisions, and 60 percent of them support proposals that would make recusal mandatory in cases that feature parties who had financially supported the presiding judge's campaign.[259]

---

[258] A more in-depth discussion of why the treatment letter had no discernable impact on judicial behavior in regard to recusals is included in a preceding, coauthored work. See Jonathan S. Krasno, Donald P. Green, Costas Panagopoulos, Michael Schwam-Baird, and Dane Thorley, *Please Recuse Yourself: A Field Experiment Exploring the Relationship Between Campaign Donations and Judicial Recusal,* working paper.

[259] Justice at Stake*, State Judges Frequency Questionnaire* (2002), available at http://www.justiceatstake.org /media/cms/JASJudgesSurveyResults_EA8838C0504A5.pdf.

The data also show that attorneys do not ask judges to recuse, even when they are aware of the donation(s) from the other attorney. Increased disclosure is one of the most common procedural solution suggested by both skeptics and supporters of the current recusal regime, but attorneys in the study were no more likely to request recusal when the judge in induced to disclose by the experimental treatment. This does not mean that disclosure does nothing to remedy judicial bias—judges who disclose may engage in higher levels of de-biasing and attorneys may be more likely to appeal the outcomes of cases in which they know the judge may have been biased—but it does indicate that increasing disclosure is not enough to lead to more requests for recusals, at least due to potential bias stemming from campaign contributions.[260]

These results should be understood in the context of the study, which involved trial-level judges and campaign donations, but are still strong evidence that the current recusal regime employed in the United States relies on faulty assumptions regarding judicial and attorney behavior and needs reform. The tailored procedural solution we have proposed more fully accounts for the often-perverse incentives at play in making recusal determinations and should therefore result in more impartial judging and increase the public's sense of judicial legitimacy. Alternatively, we might consider addressing conflicts at their genesis through institutional reform such as anonymized donations, although such an approach is practically difficult and may not be politically tenable.

More broadly, this Chapter highlights the importance of accounting for the non-legal factors that go into adjudicative decision-making when crafting rules and procedure. It also demonstrates the value that randomized field experiments—an empirical methodology often thought to be unfeasible in the court context—can provide to the analysis of procedure and policy.

[260] It is possible that attorneys do not believe that judges should recuse simply because of campaign contributions. Alternatively, they may have already known about the donation. In both cases, disclosure would have no effect on their propensity for requesting recusal for reasons unrelated to the reasons we outline in this Chapter.

Chapter 2

Randomness Pre-Considered: Recognizing and Accounting for "De-Randomizing" Events When Utilizing Random Judicial Assignment

I. Introduction

In his 2010 article titled *Randomness Reconsidered: Modeling Random Judicial Assignment in the U.S. Courts of Appeals*,[1] Matthew Hall challenged the long-standing assumption that judges in the U.S. Federal Courts of Appeals were randomly assigned to their cases.[2] Specifically, Hall argued that earlier research claiming to identify a strong causal relationship between the ideologies of judges and the tenor of their decisions[3] was systematically flawed because the investigators had assumed that all cases were assigned to circuit panels on a random basis. After demonstrating this assumption to be erroneous and accounting for the courts' actual assignment procedures (some of which were, if fact, random), Hall found that the researchers had identified ideological effects where none existed and failed to identify effects for circumstances in which ideology did indeed make a difference.

The mistakes highlighted by Hall are not uncommon in the courts and judicial behavior literature. An increasingly large number of empirical articles have utilized random assignment of judges to make causal claims regarding an array of legal, political, and economic outcomes. These studies, many of which are reviewed below, rely on the fact (or often the assumption) that the judges in their samples are assigned to cases on a truly random basis. This is because random assignment allows researchers to overcome the most important and difficult hurdle associated with causal identification—controlling for unobserved heterogeneity. Despite the essential role that random case assignment plays in such studies, however, researchers often fail to delve into the specific mechanics of a court system's assignment process, choosing instead to

[1] Matthew Hall, *Randomness Reconsidered: Modeling Random Judicial Assignment in the U.S. Courts of Appeals*, 7 J. EMP. LEGAL STUD. 574 (2010).

[2] The assumption of random assignment is so embedded in legal research that the "random assignment of federal appellate judges to panels has become a 'hallmark' of the system." Emerson H. Tiller & Frank B. Cross, *A Modest Proposal for Improving American Justice*, 99 COLUM. L. REV. 215, 216 (1999).

[3] Cass R. Sunstein, Lisa Michelle Ellman, & David Schkade, *Ideological Voting on Federal Courts of Appeals: A Preliminary Investigation*, 90 VA. L. REV. 301 (2004).

rely on the assumptions of previous studies or the simple assurances of court officials.[4] In reality, however, verifying random assignment and using the resulting data to make causal claims is often a complicated process and should mirror the multi-step analysis conducted in researcher-driven randomized studies such as field experiments.

This Chapter adds to the growing body of methodological literature exploring the use of random judicial assignment by identifying a set of common assignment procedures that I call "de-randomizing" events.[5] These events, which include non-random assignment itself, must be accounted for in order to make unbiased causal claims but are commonly either ignored or not even recognized by researchers relying on random judicial assignment. In classifying and illustrating these events, I also highlight the increasing need for legal empiricists to develop field-specific methods that integrate robust statistical techniques with the specialized procedural and doctrinal knowledge required to adequately study the law.[6]

---

[4] This is not to suggest that courts are being misleading (although some are—*see infra* note 100). Rather, the colloquial use of "random" is not equivalent to the statistical definition. *See infra* Subpart III.A.1 for a more in-depth discussion of this confusion.

[5] In addition to the Hall 2010 piece (*supra* note 1), *see* J. Robert Brown, Jr. & Allison Herren Lee, *Neutral Assignment of Judges at the Court of Appeals*, 78 TEX. L. REV. 1037 (2000); and Burton M. Atkins & William Zavoina, *Judicial Leadership on the Court of Appeals: A Probability Analysis of Panel Assignment in Race Relations Cases on the Fifth Circuit*, 18 AM. J. POL. SCI. 701 (1974).

Two recent articles also explore the panel composition procedures (as opposed to the case assignment procedures) in U.S. Circuit Courts in detail. *See* Adam S. Chilton & Marin K. Levy, *Challenging the Randomness of Panel Assignment in the Federal Courts of Appeal*, 101 CORNELL L. REV. 1 (2015); and Marin K. Levy, *Panel Assignment in the Federal Courts of Appeal*, 103 CORNELL L. REV. 65 (2017).

It is also worth noting that the popular press has expressed concerns regarding the non-randomness of panel composition and case assignment in the federal courts, although for non-methodological reasons. *See* Joe Palazzolo, *The Problem With Not-so-random Case Assignment*, WALL ST. J., November 4, 2014; Benjamin Weiser & Joseph Goldstein, *Federal Court Alters Rules on Judge Assignments*, N.Y. Times, December 23, 2013; and Alison Frankel, *Chief Judge: Rakoff Assignment to Citi Case Was "Totally Random",* REUTERS, November 30, 2011.

[6] I agree with Epstein and King's proclamation that "the law is important enough to have a subfield devoted to methodological concerns, as does almost every other discipline that conducts empirical studies." Lee Epstein & Gary King, *The Rules of Inference*, 69 U. CHI. L. REV. 1, 11 (2002). They go on to say that "[u]nfortunately, the complete list of all law review articles devoted to improving understanding explicating or adapting the rules of inference is as follows: none." *Id.* While this is certainly not quite as true anymore, the underlying sentiment is still reflective of the field.

Additionally, I attempt to fill in what others have noted[7] to be a dearth of information on the assignment protocols of courts other than the U.S. Courts of Appeals by presenting original data from a survey of the 30 largest state-level criminal courts in the United States, outlining their assignment protocols, and identifying the extent to which they feature the "de-randomizing" events highlighted in this Chapter.

This Chapter proceeds in four parts. In Part II, I review the basic criteria required for unbiased causal inference, explain the empirical advantages of random assignment, and provide a short review of the literature that has utilized random assignment of judges. In Part III, I discuss the various "de-randomizing" events that researchers should account for in order to make unbiased causal inferences and provide solutions—where they exist—for those events. In Part IV, I present original data from a survey of state-level criminal justice systems on the judicial assignment procedures used in their courts. I briefly conclude in Part V.

II. Random Assignment and Causal Inference

Over the last decade, the volume of legal scholarship utilizing random assignment has grown rapidly.[8] Researchers have randomly made offers of legal representation,[9] randomly

[7] *See* Adam M. Samaha, *Randomization in Adjudication*, 51 WM. & MARY L. REV. 1, 47 (2009).

[8] While still not predominantly featured in law reviews—due in part to their quantitative nature—the number of articles featuring randomized experiments or naturally occurring randomizations that have appeared in top law journals have dramatically increased in recent years. *See* Figure 1 in Donald P. Green & Dane R. Thorley, *Field Experimentation and the Study of Law and Policy*, 10 ANN. REV. L. & SOC. SCI. 53, 56 (2014).

[9] *See* Carroll Seron, Greg Van Ryzin, Martin Frankel, & Jean Kovath, *The Impact of Legal Counsel on Outcomes for Poor Tenants in New York City's Housing Court: Results of a Randomized Experiment*, 35 L. & SOC'Y REV. 419 (2001) (in which the authors randomly offered legal assistance to individuals involved in property disputes); D. James Greiner & Cassandra Wolos Pattanayak, *Randomized Evaluation in Legal Assistance: What Difference Does Representation (Offer and Actual Use) Make?*, 121 YALE L. J. 2118 (2012) (in which the authors randomly offered legal assistance to individuals seeking unemployment benefits); and D. James Greiner, Cassandra Wolos Pattanayak, & Jonathan Hennessy, *The Limits of Unbundled Legal Assistance: A Randomized Study in a Massachusetts District Court and Prospects for the Future*, 126 HARV. L. REV. 901 (2013) (in which the authors randomly offered legal assistance to individuals involved in landlord-tenant disputes).

manipulated jury instructions,[10] and of particular relevance to this Chapter, taken advantage of court systems that randomly assign judges to cases. By their nature, these studies are often expensive, logistically complicated, and require particular circumstances under which randomization is feasible and the resulting data are available.[11] Why then, despite these difficulties and the relatively abundant availability of non-randomized data, have we seen this increase in the use of randomization in empirical legal studies? The answer lies in randomization's ability to provide researchers with unbiased estimates of causal effects by mitigating what is arguably the most intractable obstacle associated with causal identification: controlling for unobserved heterogeneity.

In this Part, I outline the problem of unobserved heterogeneity and explain how random assignment helps researchers overcome it. In doing so, I rely on Imbens and Rubin's potential outcomes framework and accompanying notation.[12] While these concepts will undoubtedly be familiar to many and, conversely, daunting to others, an increased understanding of the fundamentals of causal inference will be valuable to all who are interested in working with randomized data. Furthermore, the intuition and notation featured in this Part form the basis for exploring the "de-randomizing" events that I discuss in Part III.

A. Unobserved Heterogeneity as a Barrier to Causal Inference

Most causal questions can be simplified down to a relationship (or set of relationships) between three elements: the causal factor or treatment, which is commonly referenced using "*X*"; an outcome of interest, "*Y*"; and the unit or subject being studied. As an example of a causal question that researchers studying courts might be interested in (an example we will return to

[10] *See* Larry Heuer & Steven D. Penrod, *Instructing Jurors: A Field Experiment with Written and Preliminary Instructions*, 13 L. & HUM. BEHAV. 409 (1989).

[11] *See* ALAN S. GERBER & DONALD P. GREEN, FIELD EXPERIMENTS: DESIGN, ANALYSIS, AND INTERPRETATION 13-15 (2012).

[12] *See* Donald B. Rubin, *Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies*, 6 J. EDU. PSY. 688 (1974). *See also* Chapters 1 and 2 in GUIDO W. IMBENS & DONALD B. RUBIN, CAUSAL INFERENCE FOR STATISTICS, SOCIAL, AND BIOMEDICAL SCIENCES: AN INTRODUCTION (2015).

frequently in this Chapter), let us suppose that we want to know the impact of a criminal case (our subject) being assigned to either a female or male judge (our *X*) on the length of defendant's sentence (our *Y*).[13] As a first step in answering this question, we might gather data on cases—including outcomes in regard to the length of final sentences—and compare the cases in which the presiding judge was female to those in which the judge was male. As it turns out (in our hypothetical study), female judges tend to impose longer sentences than male judges. Should we then conclude that being assigned to a female judge results in longer criminal sentences?

Many will object to this basic analysis, correctly observing that the cases assigned to female judges might be systematically different than the cases assigned to male judges in a number of important ways apart from the gender of the judge. For example, female judges in our hypothetical court system might be more likely to preside over sexual assault cases, which generally feature longer average sentences than other types of cases. Or maybe male judges tend to preside over geographical areas that contain wealthier residents, so defendants in their courtrooms will have higher quality attorneys and, consequently, shorter sentences.

These both seem like plausible concerns, so we might attempt to account for these confounding relationships by "controlling" for them in our statistical analysis.[14] After doing so, our original results stand: female judges seem to be more punitive than male judges. But what if it turns out that the court clerk, who is in charge of case assignment, tends to give the "toughest" cases to the more experienced judges, who happen to be female. Might we believe that case complexity has any relationship to final verdict? Probably. So we also control for case

[13] The role of a judge's gender in case outcomes has been explored in a number of studies. *See, e.g.,* Christina L. Boyd, Lee Epstein, & Andrew D. Martin, *Untangling the Causal Effects of Sex on Judging*, 52 AM. J. POL. SCI. 389 (2010); and Jennifer L. Peresie, *Female Judges Matter: Gender and Collegial Decisionmaking in the Federal Appellate Courts,* 114 YALE L. J. 1759 (2005).

[14] In observational (non-randomized) studies, researchers control for confounding variables by holding them constant when calculating the statistical association between "X" and "Y."

complexity, and our subsequent results indicate that defendants still receive more lenient sentences under male judges.

Have we identified the causal impact of being assigned a male judge? Maybe, but we cannot be sure. Despite our repeated attempts at controlling for possible sources of heterogeneity between judges such as case type, jurisdiction, and experience, there might be some other factor driving the relationship between our *X* and our *Y* that we have not considered. Guilty defendants may, for example, be more likely to make peremptory challenges or requests for judicial transfer when assigned to male judges, or maybe female judges are assigned to hear cases right before lunchtime and are therefore, more prone to foul moods during sentencing proceedings.[15] Ultimately, the relationship between the gender of the judge and the length of a criminal sentence is deceptively complex, and controlling for all the factors in that relationship through *post-hoc* analysis is extremely difficult, if not impossible.

The issue we keep running into is the result of two significant barriers to causal identification: First, the laws of nature prevent us from comparing a given case assigned to a male judge against that same exact case assigned to a female judge. This inherent limitation, titled the "fundamental problem of causal inference" by Paul Holland, intuitively posits that you can never simultaneously observe the effect of more than one treatment on the same subject and therefore, cannot ever truly identify the causal impact of a given factor.[16]

In the process of doing our best to overcome the fundamental problem of causal inference in the discussion above, we attempted to construct two similarly balanced comparison groups and subsequently encountered the second crucial barrier to unbiased causal inference:

---

[15] As a good example of the problems this Chapter is attempting to highlight, the "lunch effect" that was identified in a 2011 paper (Shai Danzinger, Jonathan Levav, & Liora Avnaim-Pesso, *Extraneous Factors in Judicial Decisions*, 108 PNAS 6889 (2011)) has subsequently been criticized—by the court system that the original study evaluated—for assuming random assignment when the order of case assignment was largely determined by whether the prisoner was represented by an attorney. Karen Weinshall-Margel & John Shapard, *Overlooked Factors in the Analysis of Parole Decisions*, 108 PNAS (Letter) E883 (2011).

[16] Paul W. Holland, *Statistics & Causal Inference,* 81 J. AM. STAT. ASS'N. 945 (1986).

unobserved heterogeneity. Try as we might, it is extremely difficult to create artificial comparison groups by controlling for variables that we believe might be interfering with our X-causes-Y inference. Even more difficult—likely impossibly difficult—is knowing whether we have identified all of the confounding variables influencing the outcomes in our cases. Causal relationships, even those with which we are familiar, are deviously complex, making it difficult for us to be confident in statistical methods that rely on the assumption that all the important variables have been correctly modeled.

B. Random Assignment as a Solution to Unobserved Heterogeneity

How does random assignment help us overcome these two key problems? Clearly it does not solve the fundamental problem of causal inference (we still observe only one outcome for each case), but it does allow us to remedy the problem of unobserved heterogeneity by creating comparison groups that, in expectation, are statistically indistinguishable from each other in every way, even those ways that we are not aware of or able to observe.

In every cause and effect relationship, we can identify a treatment and at least one outcome.[17] Let us use the indicator "d_i" to designate the type of treatment that an individual is assigned to, where the subscript "i" symbolizes any given individual in the study sample. At this point, the treatment mechanism can be random assignment or some natural, non-random assignment process. d_i will take on a value of 1 if subject i is treated and 0 if that subject is not treated. Continuing with the example used above—the effect of being assigned a male judge on sentencing lengths—d_i would equal 1 if the assigned judge is male and 0 if the judge is female.[18]

[17] For the sake of simplicity, this Chapter will generally use "treatment" to refer to one of the possible assignment groups, and "control" to refer to the other assignment group even though many research questions do not technically have a "control" group.

[18] Considering gender is generally a fixed characteristic, it may be odd to think of it as a "treatment." However, the gender of the judge that is assigned to a given case can vary because the judicial assignment itself can vary. Also note that this hypothetical research question only gets to the effect of being assigned a male judge, which is not necessarily equivalent to the effect of the judge being male (male judges may have specific characteristics that are correlated only with gender as opposed to being caused by gender).

Now imagine that instead of outcomes being denoted simply using "Y", every individual in the dataset has two *potential* outcomes, $Y_i(d_i = 1)$ and $Y_i(d_i = 0)$, where $Y_i(d_i = 1)$, or $Y_i(1)$ for the sake of simplicity, represents the outcome if an individual is treated (e.g. the length of sentence if an individual is assigned to a male judge) and $Y_i(0)$ represents the outcome if an individual is not treated. If it seems counterintuitive for an individual to have two outcomes, remember that these are *potential* outcomes, meaning that neither outcome has been realized. We cannot observe both potential outcomes at the same time, but if we could, it would follow that the difference between them, $Y_i(1) - Y_i(0)$, would equal the treatment effect of being assigned a male judge (relative to a female judge) on the final verdict or outcome for individual i.

Applying this principle to all the individual cases in the dataset, we can define the average or *expected* treated potential outcome for the entire sample as $E[Y_i(1)]$ (where $E[-]$ is the expected value) and the expected untreated potential outcome as $E[Y_i(0)]$. Taking the difference between these two values, $E[Y_i(1)] - E[Y_i(0)]$, gives us the average treatment effect, or ATE, of the entire sample of cases. However, since we cannot observe both potential outcomes for any of the individuals, the best we can do is take the average outcomes for those cases that we observe in the treatment group and compare them to the average outcomes for those we observe in the control group:

$$E[Y_i(1)|d_i = 1] - E[Y_i(0)|d_i = 0] = \widehat{\text{ATE}}.^{19}$$

As we discussed above, however, this estimator of the $\widehat{\text{ATE}}$ is likely biased because the two expected values are probably different in some important way other than the treatment condition because the mechanism through which an individual is assigned a treatment condition, d_i, may be related to some confounding variable (observed and/or unobserved, we cannot know for sure). We did not see how and why a given i received a $d_i = 1$ or $d_i = 0$, so

[19] $\widehat{\text{ATE}}$, or ATE-hat, is shorthand for the estimate of the ATE.

encountering the problem of unobserved heterogeneity is inevitable, thus preventing us from trusting the resulting estimate of the ATE.

Random assignment of treatment conditions overcomes unobserved heterogeneity by assigning individuals to treatment groups independent of any other factor. In other words, the treatment condition to which an individual is assigned is not related to type of crime, seniority of judge, or any other possibly important pre-treatment characteristic, including the ones we do not know about or are unable to measure. Importantly, this also means that an individual's d_i is assigned independently of an individual's potential outcomes, so even though we cannot observe each group's full set of potential outcomes, we know that, in expectation, the potential outcomes for each comparison group will be identical. Therefore, the expected $Y_i(1)$'s for those assigned to the treatment group, $E[Y_i(1)|d_i = 1]$, will be identical to the expected $Y_i(1)$'s for those in the control group, had they been assigned to the treatment group, $E[Y_i(1)|d_i = 0]$. The same is true of each group's expected $Y_i(0)$'s, enabling us to identify an accurate measure of the overall expected value of both potential outcomes without actually observing both outcomes for a given individual:

$$E[Y_i(1)|d_i = 1] = E[Y_i(1)|d_i = 0] = E[Y_i(1)] \text{ and}$$

$$E[Y_i(0)|d_i = 1] = E[Y_i(0)|d_i = 0] = E[Y_i(0)].$$

This, in turn, allows us to take the difference between the actual outcomes for the two treatment groups and calculate an unbiased estimate of the ATE:

$$E[Y(1)] - E[Y(0)] = E[Y(1) - Y(0)] = \widehat{\text{ATE}}.$$

When the assignment to treatment groups is truly random, the resulting $\widehat{\text{ATE}}$ is an unbiased estimate of the effect of the treatment on the outcome of interest, overcoming the potential bias created by unobserved heterogeneity and circumventing the restrictions highlighted by the fundamental problem of causal inference. When properly utilized by

researchers,[20] either as an implemented strategy or, in the case of randomly assigned judges, as an existing attribute of pre-existing, observational data, random assignment allows researchers to answer important causal questions in a reliable manner.

C. How Researchers Have Utilized Random Assignment of Judges

Realizing the potential of random assignment to answer important causal questions, scientists have long seen randomized experiments as the gold standard in research design. While research harnessing random assignment is not as common in legal studies (including law-oriented political science and economics) as it is in the "hard" sciences or even the social sciences generally,[21] randomization's use as a tool for studying legal questions has seen marked growth in the last decade.[22] Studies utilizing the random assignment of judges are no exception to this trend. The researchers authoring these pieces have addressed a range of interesting and important topics, and while many them do not address the methodological concerns highlighted in this Chapter, their analyses nonetheless provide us with valuable examples of what can be accomplished.

Not surprisingly, a large number of these researchers are interested in exploring the role of a judge's gender, race, and ethnicity in their legal decisions.[23] In Abrams et al.'s substantial

---

[20] There are a number of additional and important assumptions that are required when utilizing random assignment that are not discussed in this Chapter, including excludability (often called the exclusion restriction) and non-interference (often called the Stable Unit Treatment Value Assumption, or SUTVA). For an in-depth analysis of why these assumptions are required and how they should be addressed, *see* IMBENS & RUBIN, *supra* note 12; and GERBER & GREEN, *supra* note 11.

[21] For a review of the growth of randomized experiments in political science, *see* James N. Druckman, Donald P. Green, James H. Kuklinski, & Arthur Lupia, *Experimentation in Political Science* in CAMBRIDGE HANDBOOK OF EXPERIMENTAL POLITICAL SCIENCE 3-14 (James N. Druckman, Donald P. Green, James H. Kuklinski, and Arthur Lupia, eds., 2010). For the same in economics, *see* Steven D. Levitt & John A. List, *Field Experiments in Economics: The Past, Present, and Future*, 53 EURO. ECON. REV. 1 (2009).

[22] *See* Figure 1 in Green & Thorley, *supra* note 8.

[23] It is important to note that the appropriate estimand (that which is estimated) in these studies is the *effect of being assigned* a judge with a particular characteristic (race/gender/ideology) as opposed to the *effect of that characteristic*. Each judge has a bundle of characteristics, and without randomly assigning those individual characteristics (unfeasible for race, gender, and ideology), we are unable to identify the source of a causal effect. By randomly assigning the judge, however, we are able to identify the causal impact of assignment to a judge with a particular characteristic.

analysis of the Cook County (Chicago) Criminal Court, the authors provide evidence of significant racial disparities in the rates that black and white judges incarcerate black and white defendants, although these disparities disappear in regard to the sentence lengths that incarcerated defendants receive.[24] Similarly, Gazal-Ayal and Sulitzeanu-Kenan find that Arab and Jewish judges in Israeli bail hearings hand down more lenient detention decisions to defendants in their ethnic in-group.[25] Gender also appears to be an important factor in judicial decision-making. Eisenberg et al.'s 2012 study on Israeli courts finds that female judges are as much as 15% more likely to cast a judicial vote in favor of defendants.[26] More recent research indicates the influence of gender may extend beyond just the decision in a given case. Farhang et al. utilize random panel assignment in the U.S. Courts of Appeals to show that female-majority panels are more likely to write opinions in sexual harassment cases, thereby being more likely to have substantial influence on the overall jurisprudence in that area.[27]

Studies have also looked at the causal effects of being assigned to a judge with a particular political ideology or set of beliefs. Focusing almost exclusively on U.S. federal courts,[28] authors have compared the impact of being assigned to conservative and liberal judges (generally defined using the political party of the president who appointed them) on an array of judicial outcomes. The results have been mixed.[29] Beim and Kastellec test the effect of the

---

[24] David S. Abrams, Marianne Bertrand, & Sendhil Mullainathan, *Do Judges Vary In Their Treatment of Race?*, 41 J. LEGAL STUD. 347 (2012).

[25] Oren Gazal-Ayal & Raanan Sulitzeanu-Kenan, *Let My People Go: Ethnic In-group Bias in Judicial Decisions—Evidence from a Randomized Natural Experiment*, 7 J. EMP. LEGAL STUD. 403 (2010).

[26] Theodore Eisenberg, Talia Fisher, & Issi Rosen-Zvi, *Does the Judge Matter? Exploiting Random Assignment on a Court of Last Resort to Assess Judge and Case Selection Effects*, 9 J. OF EMP. LEGAL STUD. 246 (2012).

[27] Sean Farhang, Jonathan P. Kastellec, & Gregory J. Wawro, *The Politics of Opinion Assignment and Authorship on the U.S. Court of Appeals: Evidence from Sexual Harassment Cases*, 44 J. LEGAL STUD. 59 (2015).

[28] As I discuss at the end of this Part, there are a number of explanations for the exclusive focus on federal courts of appeals, but the research focused on identifying the impact of ideology on judicial decision-making can—and, in fact, should—be expanded to state-level court systems.

[29] This may be due to a difference in outcome measures but is also possibly due to similar mistakes that Hall addressed in his 2010 article (see *supra* note 1) and that I am highlighting in this Chapter.

ideological composition of three-judge panels in the U.S. Courts of Appeals on decisions regarding death penalty and find that the ideologies of the judges has a significant impact on the outcome of the case.[30] Going further, they also show that ideological composition partially determines the propensity for dissent, which, as a result, affects the likelihood that a case will be heard *en banc*. Likewise, Niblett and Yoon's study on U.S. Courts of Appeals suggests that ideology has a significant relationship with the way in which legal precedent is used.[31] For a number of case types, conservative panels are much more likely to cite conservative precedent and overturn liberal precedent, with the opposite being true for liberal panels. Conversely, Ashenfelter et al. find no ideological effect on case outcomes,[32] and, as was mentioned at the beginning of this Chapter, Hall's research has suggested that the effect of ideological composition depends on the particular court and the content of the case.[33]

Researchers have also discovered ways to combine the benefits of randomly assigned judges with other empirical methods to expand the realm of issues that can be explored. For example, a number of recent studies have used instrumental variable design[34] to identify the downstream impacts that incarceration has on the lives of criminal defendants and their families. While it would be impractical—not to mention possibly unethical[35]—to intentionally randomize the amount of time that an individual spends in jail, these researchers have taken

---

[30] Deborah Beim & Jonathan P. Kastellec, *The Interplay of Ideological Diversity, Dissents, and Discretionary Review in the Judicial Hierarchy: Evidence from Death Penalty Cases*, 76 J. POL. 1074 (2014).

[31] Anthony Niblett & Albert H. Yoon, *Friendly Precedent*, 57 WM. & MARY L. REV. 1789 (2016).

[32] Orley Ashenfelter, Theodore Eisenberg, & Stuart J. Schwab, *Politics and the Judiciary: The Influence of Judicial Background on Case Outcomes*, 24 J. LEGAL STUD. 257 (1995).

[33] Hall, *supra* note 1.

[34] For a more in-depth explanation of how these models can be used to identify downstream effects, see Chapter 6 in GERBER & GREEN, *supra* note 11.

[35] While it may initially seem unethical to randomly assign something as serious as sentence lengths, many (including the author of this Chapter) argue that laws and legal procedures are no more immune to randomization than medical treatments and pharmaceuticals. For a particularly sophisticated discussion of this topic, *see, e.g.,* Michael Abramowitz, Ian Ayres, & Yair Listokin, *Randomizing Law*, 159 U. PA. L. REV. 929 (2011).

advantage of the natural disparities in punitiveness that exists between judges[36] combined with the random assignment of those judges to find an unbiased estimate of the effects of incarceration on a number of important outcomes. Green and Winik find that increased time spent in jail has no distinguishable impact on the likelihood that an individual will recidivate,[37] although a similar study by Loeffler comes to the opposite result.[38] That same study by Loeffler also suggests that increased prison time leads to higher post-sentence unemployment, a finding that is supported by Kling's 2006 study on the same topic.[39] These approaches have recently been expanded upon to measure the impact that incarceration has on juvenile offenders[40] and the broader labor market and economy.[41] A group of recent papers have similarly used random or as-if random judicial assignment as an instrument for the assignment of bail and pretrial detainment in criminal cases, providing novel empirical estimations for the effect of detainment on recidivism, case outcomes, and future employment.[42]

---

[36] While identifying inter-judge disparity is not necessarily a causal question, it is a question that can only reliably be answered by analyzing a system with random assignment because otherwise, difference between judges may be the result of unobserved heterogeneity resulting from the assignment process. *See* James M. Anderson, Jeffrey R. Kling, & Kate Stith, *Measuring Inter-Judge Sentencing Disparity Before and After the Federal Sentencing Guidelines*, 42 J. L. & ECON. 271, (1999).

[37] Donald P. Green & Daniel Winik, *Using Random Judge Assignments to Estimate the Effects of Incarceration and Probation on Recidivism Among Drug Offenders*, 48 CRIMINOLOGY 357 (2010).

[38] Charles E. Loeffler, *Does Imprisonment Alter the Life Course? Evidence on Crime and Unemployment from a Natural Experiment*, 51 CRIMINOLOGY 137 (2013).

[39] Jeffrey R. Kling, *Incarceration Length, Employment, and Earnings*, 96 AM. ECON. REV. 863 (2006).

[40] *See* Anna Aizer & Joseph J. Doyle, Jr., *Juvenile Incarceration, Human Capital and Future Crime: Evidence from Randomly-Assigned Judges*, 130 Q. J. ECON. 759 (2015).

[41] *See* Michael Mueller-Smith, *The Criminal and Labor Market Impacts of Incarceration*, Working Paper (2015) (which also identifies the individual impact of incarceration on recidivism, future employment, and the likelihood of marriage).

[42] Arpit Gupta, Christopher Hansman, & Ethan Frenchman, *The Heavy Costs of High Bail: Evidence from Judge Randomization*, 45 J. LEGAL STUD. 471 (2016); Emily Leslie & Nolan G. Pope, *The Unintended Impact of Pretrial Detention on Case Outcomes: Evidence from New York City Arraignments*, 60 J. L. & ECON. 529 (2017); Will Dobbie, Jacob Goldin, & Crystal S. Yang, *The Effects of Pretrial Detention on Conviction, Future Crime, and Employment: Evidence from Randomly Assigned Judges*, 108 AM. ECON. REV. 201 (2018); and Megan Stevenson, *Distortion of Justice: How the Inability to Pay Bail Affects Case Outcomes,* 34 J. L. Econ. & Org. 511 (2018).

In a particularly novel utilization of this methodology, Dahl et al. conducted a study that uses random judicial assignment to gauge the existence and extent of "welfare cultures," or the propensity for a parent's participation in welfare programs to increase the likelihood that children will do so as well.[43] Because applicants for Norwegian disability insurance appeal their decisions to judges who are randomly assigned, Dahl et al. were able to show that parental participation in welfare leads to a 6-12 percentage point increase in the probability that children will also be participants as adults. Leibovitch also provided a fresh example of this methodological approach by exploiting random assignment to identify the impact that the cases a judge presides over early in her career has on her future decisions, finding that judges who are initially exposed to less serious criminal cases are more punitive later in their careers.[44]

While these studies provide an illustration of the diversity of issues that can be investigated using randomly assigned judges, the methodological strategy is relatively young and is almost certainly underutilized by researchers interested in making causal inferences in the court context. It is notable, for example, that nearly half of the articles listed above take place in the context of the U.S. Federal Courts of Appeals—a court venue that features only a small percentage of all the legal cases in the United States.[45] Going forward, researchers are sure to find more ways to utilize its benefits, although the extent to which these studies will accurately inform our understandings of law, policy, and economics will be largely dependent on whether researchers understand and account for the large number of ways in which a seemingly random assignment protocol is, in fact, not random.

---

[43] Gordon B. Dahl, Andreas Ravndal, & Magne Mogstad, *Family Welfare Cultures*, Q. J. ECON. 1 (2014).

[44] Adi Leibovitch, *Relative Judgments*, 45 J. LEGAL STUD. 281 (2016).

[45] There are a number of possible reasons for this imbalance: Empirical legal scholars, especially political scientists, tend to be more substantively interested in the federal courts of appeals than with federal district courts and state courts. Additionally, and of particular relevance to this Chapter, researchers may perceive state and district courts as less procedurally amenable to the assumptions required to utilize random judicial assignment.

III. "De-Randomizing" Events in Judicial Assignment Procedure

The articles reviewed above make the same assumption regarding the assignment of judges in their datasets—that it was random, ostensibly allowing researchers to overcome the problem of unobserved heterogeneity. It is not clear, however, that all of them investigated the assignment procedures closely enough to warrant such conclusions. Over the life of a case, the process involved in determining who will be the presiding judicial authority is deceptively nuanced. Some of this nuance is not problematic in regard to making unbiased causal claims, but much of it is a threat to random assignment and must therefore be accounted for in order to determine whether a dataset can yield reliable estimates of treatment effects.

Although studies utilizing random judicial assignment are relying on the same basic estimation strategy as researcher-driven randomized studies such as field and lab experiments, they often do not meet (or at least demonstrate or discuss) the same basic assumptions that any researcher-driven experiment would be expected to address. A methodological review of any formal experiment would, for example, look for instances of assignment error, non-compliance to treatment conditions, stable unit treatment value violations (i.e. spillover), and attrition. Researchers should assume that naturally-occurring randomization is just a messy as researcher-driven experimentation. In fact, given that the randomization schemes used in courts are generally not designed to facilitate empirical studies but instead to equitably assign cases and avoid the appearance of impropriety, it is reasonable to expect even more systematic threats to inference.

Conversely (and more positively), properly understanding the interaction between court procedure and causal identification will open the door to studying venues and substantive topics that would otherwise be overlooked. As researchers in law, economics, political science, and other related fields begin to exploit the advantages of random judicial assignment with greater

frequency, some may conclude—generally through statistical balance tests[46]—that datasets are seemingly non-random and therefore unusable when they are actually ripe for meaningful empirical analysis. As I explain below, however, many of the non-random elements in and otherwise random judicial assignment scheme can be accounted for in ways that allow unbiased causal inference.

In this Part, I explore the various ways in which a jurisdiction's assignment procedure may violate the assumption of random assignment by outlining four key "de-randomizing" events:[47] 1) the initial assignment procedure, 2) differing probability of assignment, 3) post-assignment judicial changes, and 4) missing outcomes. While each of these general events should be familiar to those versed in experimental analysis, diagnosing the specific manifestations of them in the judicial assignment context requires a solid understanding of assignment procedure. I begin each subpart with an explanation of the threat the event poses to unbiased causal inference, after which I provide examples of the most common ways in which these events occur within judicial assignment procedure and outline how they might be addressed. I also provide a brief overview of how researchers can verify that they have accurately modeled these events using quantitative and qualitative tests.

While these events are explicated in the context of trial courts, they are generally applicable to appellate level assignment procedures as well. It should also be noted that this list of events should serve only as a starting point for researchers and readers working with random judicial assignment; not every court system will feature all of these events, and the list is not comprehensive.

[46] *See infra* Subpart III.E for a detailed discussion of these tests and their limitations.

[47] It does not appear that "de-randomizing" is a term of art in any discipline. However, it is a helpful term for encompassing all of the events that are potential threats to the benefits provided by random judicial assignment.

A. The Initial Assignment Procedure

The Problem: The first and most obvious way in which a court's assignment procedure might introduce bias is if the basic assignment procedure is not random. The idea that the utilization of randomly assigned judges requires random assignment may seem tautological, but the distinctions between random and "not-quite-random" procedure are not always clear. Additionally, even in a system that normally assigns judges to cases in a truly random way, institutionalized exceptions may introduce significant levels of bias by essentially "de-randomizing" a portion of the available cases. In this Subpart I categorize three types of judicial assignment procedures, discuss their usability as mechanisms for making unbiased causal inferences, and highlight three common non-random exceptions to the regular (often random) assignment process: case specialization, administrative discretion, and assignment based on a party's previous interaction with the court.

i. Different Types of Judicial Assignment: The procedures through which courts assign cases to judges can be broadly categorized into three varieties: random assignment, in which assignment is done according to a statistically random process; as-if random assignment, in which assignment is not done randomly but is done in a way that, under a number of assumptions, can still allow for unbiased causal identification; and non-random assignment, in which cases are assigned in a way that prevents researchers from estimating unbiased treatment effects. Understanding the differences between these three categories is vital, although recognizing them is often problematic; many court systems have a mixed set of procedures (part random and part non-random), and court administrators may even state that the procedures they use are random, when they are in fact as-if random or even completely non-random.[48]

[48] This occurred a large number of times over the course of the court survey I discuss in Part IV, *infra*. Upon asking them to describe the basic assignment procedures used in their court, administrators (judges, clerks, etc.) would often respond by telling me that judges were assigned to cases on a "random basis," or by "completely random selection." After asking a number of follow up questions, however, it became clear that what they meant by "random" would more accurately be described as "arbitrary." Cases were not, for example, assigned by a random-number generator or by pulling names out of a hat (both viable random

Statistically random judicial assignment can be defined as an assignment procedure in which cases are allocated independently of any value, characteristic, or variable other than an exogenous assignment mechanism. Using the potential outcome notation presented above, this means that a case's treatment category ($d_i = 1$ or $d_i = 0$) should be completely unrelated to *any* of that case's pre-treatment characteristics such as case type, complexity, or the types of individuals involved in the case. The specific assignment mechanism used by a court can be any process under which the probability of assignment is known by the court (or the researcher) and is greater than 0 and less than 1 (meaning that any given case has at least some chance of being assigned to each treatment category). Such assignment mechanisms are most commonly computer-generated random numbers but may be as rudimentary as sequentially drawing judges' names from an envelope (the method for assigning homicide cases in the Nassau County Criminal Court) or a hat.

As-if random assignment, on the other hand, occurs when a case's treatment category (d_i) *is* based on one or more of that case's pre-treatment characteristics (i.e. assignment is not random) but, importantly, pre-treatment characteristics that are unrelated to that individual's potential outcomes ($Y_i(1 \text{ or } 0)$). If those pre-treatment characteristics are truly exogenous (a strong assumption), then the as-if random assignment mechanism functions in the same way as a set of random numbers, allowing the researcher to derive an unbiased estimate of the ATE for the outcome measure she is interested in. As with random assignment, as-if random assignment must feature a probability of assignment that is both between 0 and 1 and knowable by the researchers, which will require the researchers to collect data on whatever pre-treatment characteristic the assignment procedure is based on. Common as-if random assignment mechanisms include cycling lists (commonly referred to as "wheels") and schedules, which assign cases based on the order in which they come in. Despite the added assumptions that

assignment procedures), but instead according to the day of the week, the last name of the defendant, or a reoccurring list.

these types of systems necessitate, studies utilizing as-if random assignment are common, likely because as-if random assignment is used in many courts.

The final category of assignment procedure, non-random assignment, occurs when an individual's treatment category is based on one or more pre-treatment characteristics that are related to an individual's potential outcomes. Common examples of clearly non-random assignment might include assignments based off of the type of crime committed, the criminal history of the defendant, or the geographic area in which the crime was committed. These types of systems cannot reliably be used to derive unbiased causal effects without running into the unknown heterogeneity bias that was discussed above.

The distinctions between non-random and as-if random assignment are not always well defined. Each of the non-random assignment protocols listed above (assignment based on crime, defendant, or area) may, under strong assumptions and with certain outcomes of interest, be considered as-if random assignment (e.g. if we believe that the type of crime committed has no impact on the length of the criminal sentence). Likewise, protocols that use seemingly exogenous variables to make assignments, such as the last name of the defendant or the day the case files are submitted, may actually be non-random (e.g. the last name of the defendant might somehow have a relationship to that defendant's sentence length).[49]

ii. Specialization: In nearly every court surveyed for this Chapter (even those in which the general assignment process is random), a number of judges specialize in certain types of criminal cases. In some courts, this specialization is formalized and built into the regular protocol—the Nassau County Criminal court, for example, has specific "sections" (each with one or more judges) for sex offences, domestic violence, certain drug violations, and DWI cases. All

[49] On an anecdotal note, I remember sitting in a courthouse foyer with an attorney I worked for as we watched the assignment board (a screen that displayed which judge the cases were assigned to), waiting for the right moment to jump in line so our case would be assigned to a judge my boss perceived to be particularly favorable—a practice I have heard is not uncommon. Whether such efforts are successful in "gaming" the assignment process or not, the experience serves as a reminder that enterprising attorneys often find ways to "de-randomize" seemingly random assignment procedures.

the cases that fall into one of those categories bypass the regular random assignment protocol and go straight to the corresponding section. In other courts, the specialization process is more unofficial, as with capital cases going to the "more experienced" judges, a decision that is usually up to the chief judge or an upper-level clerk.

Many courts, particularly federal district courts, employ an institutionalized form of specialization known as the "related case rule." This practice—while not a rule, *per se*—encourages judges or judge panels to take many or all cases that deal with a particular substantive legal issue if they have previously ruled on a case with that issue. Macfarlane chronicles, for example, the repeated use of the related case rule in the Southern District of New York and its subsequent impact on the jurisprudence relating to stop-and-frisk policies.[50]

These divergences from the general assignment process are not necessarily threats to unbiased causal inference as long as the assignments within specialized groups are also random (although this conditional assignment procedure and any impacts it has on assignment probability should be accurately modeled—see the solutions subpart below and the discussion on probabilities of assignment in Subpart III.B). Conditional assignment that is non-random, however, can be problematic in a number of ways. First, because case specialization inherently assigns judges according to a characteristic of the case, it violates the primary assumption of random assignment, which requires that treatment assignment be independent from all pre-treatment characteristics. Additionally, the characteristic in question is almost certainly related to an individual case's potential outcomes, meaning that these assignments likely fall within the "non-random" category described above.

iii. Administrative Discretion: Administrative discretion functions similarly to specialization, except deviations from the normal assignment procedure are generally done on a case-by-case basis. A presiding judge may, for example, choose to override the normal

[50] *See* Katherine A. Macfarlane, *How the Southern District of New York's 'Related Cases' Rule Shaped Stop-and-Frisk Rulings*, 19 MICH. J. RACE & L. 199 (2014).

procedure by assigning a particularly prominent or complicated case to a more experienced judge. Or a judge who is dealing with stressful personal matters might temporarily not be assigned complex and long-lasting cases, such as those dealing with capital crimes. In any case, it is highly probable that the reasons the administrative exception is made are related in important ways to the potential outcomes for the cases.

Unlike case specialization, which is generally a formalized policy, administrative discretion is utilized on an ad hoc basis, which makes it more difficult to properly account for. Exceptions to the normal assignment procedure due to administrative oversight are not always recorded in the court records, and when they are, the records may be kept in a different location than the rest of the data. This potential for bias is mitigated to some degree by the fact that these types of administrative interference do not seem to be common (infrequent deviations have smaller relative impacts on the point estimate of the ATE) but is nonetheless potentially significant and should be identified and accounted for in any study that utilizes random judicial assignment.

iv. Party's Past Participation in Court Proceedings (Repeat and Probationary Offenders and Appeals After Remand): In a surprisingly high number of criminal cases at the state level, the defendant has previously been convicted of a crime in the jurisdiction in which her or his new case is being processed.[51] Similarly, the individual accused of the crime may currently be a defendant in another, ongoing criminal case in that jurisdiction. In either circumstance, it is common for courts to re-assign these defendants' cases to the judges who previously oversaw or who are currently presiding over previous cases. This assignment may happen as a part of the initial assignment procedure or, possibly because this connection was initially overlooked, after assignment to another judge has been made (the latter of which may

[51] A 2002 report by the U.S. Bureau of Justice indicates that recidivism rates for certain type of crimes is as high as 78.8% (motor vehicle thieves), 74.6% (larcenists), and 74.0% (burglars). Patrick Langan & David J. Levin, *Recidivism of Prisoners Released in 1994*, Bureau of Justice Statistics Special Report NCJ 193427 (2002).

have implications for the probability of assignment and post-assignment changes, see Subpart III.B).

The civil parallel to these assignment protocols is re-assigning extensions or re-submissions of past civil cases to the judge or judges that oversaw the previous cases. The assignment clerk in federal circuit courts, for example, often offers to re-assign cases that have come back to the court on remand back to the panel of judges who heard the original case.[52] While one court archivist interviewed for this study[53] believed that most panels decline re-assignment (they suggested this was because the individuals who truly knew the details of the case—the law clerks who worked on the original case—are gone), he suggested that there is often pressure to take back particularly messy cases.[54]

At first glance, this type of assignment procedure may not seem problematic from a causal inference perspective. If the assignment procedure for the original case was random, then the subsequent cases returning to that same judge can be seen as extensions of that original assignment process. This conclusion, however, makes an important assumption. In order to include repeat offenders who are re-assigned to their original judges in an analysis, we have to assume that each judge in the analysis group creates the same number and type of repeat offenders. Unfortunately, research has shown that the type of judge that an individual is assigned to may have a significant impact on whether that individual will show up in court again.[55] More punitive judges, for example, tend to generate defendants with higher recidivism rates, which—unless those defendants are always committing their secondary crimes in other jurisdictions—means that some judges will see repeat offenders or probation violators at higher

[52] This is in addition to the instances where the remanding court has explicitly specified that it retains jurisdiction post-remand.

[53] For confidentiality, I am not disclosing the name of this archivist or the circuit in which they worked.

[54] Apparently, one circuit judge "rails against the refusal to take 'come-backers'" because "we helped create this mess, so we need to clean it up."

[55] *See* Green & Winik, *supra* note 37 (who use the variation in judicial punitiveness to measure the downstream effect of judicial assignment on recidivism rates).

rates than others. If we believe that the potential outcomes of cases featuring repeat offenders may differ in an important way from those featuring first-time offenders, then reassignment of repeat offenders is a likely source of bias.

The Possible Solutions to Non-Random Assignment: The first step that a researcher should take in order to avoid the pitfalls of non-random assignment is to ensure that the venue of the study is one in which cases are randomly or as-if randomly assigned. This is best done by speaking with—or even better, visiting with—the individual or individuals who are in charge of the case assignment process in the court system of interest. In Subpart IV.C of this Chapter, I provide some insights into this process of inquiry gained after speaking with 30 state-level criminal courts about their assignment procedures. While previous studies (including this one) can be used to identify potentially workable court venues, researchers should personally confirm such conclusions before beginning the often-arduous data-collection process.

When the researcher is reasonably confident that she has identified a randomized or as-if randomized assignment scheme, she should begin the process of modeling the assignment procedures to the dataset of cases. By doing this, not only can the researcher verify the information that was communicated by the court representatives, she can lay the groundwork for identifying and addressing the additional and more complex "de-randomization" events described below. Additionally, a correctly modeled assignment process opens the door for more prognostic significance tests such as Fisher's Exact Test.[56]

Even for a court system that randomly assigns cases, this modeling process will almost certainly involve a series of conditional assignments based on a case's pre-treatment covariates,

[56] Also known as randomization inference, Fisher's Exact Test makes no distributional assumptions regarding the outcome data. By simulating all possible random assignments (a process that is not possible unless the underlying assignment process is accurately modeled), Fisher's Exact test creates an exact representation of the sampling distribution of the estimated ATE—or any other estimand—and derives a p-value based on the observed outcome's relative location in that distribution. *See* R.A. FISHER, THE DESIGN OF EXPERIMENTS (1935) (for Fisher's original use of this technique); and Jake Bowers & Costas Panagopoulos, *Fisher's Randomization Mode of Statistical Inference*, Working Paper (2011) (for a more modern application).

including specialization based on case type, re-assignment of repeat and probationary offenders, and occasional administrative discretion. While many of these conditional assignments will be easy to identify and accurately model (specialization, for example should be a formal part of the court's assignment process and can be modeled conditional on case type), there will likely be some idiosyncratic conditionality such as administrative discretion that can only be identified through court records.

If the cases in any of these conditional categories are assigned non-randomly, the most prudent approach is to omit such cases from the final statistical analysis. Doing so will naturally limit the applicability of a researcher's findings (the estimated ATE may, for example, apply only to non-drug and non-sexual violence cases), but those findings will maintain the non-biased benefits of random assignment. If, on the other hand, judges within specialized sections (assuming there is more than one judge in a section) are also randomly assigned,[57] researchers can still derive unbiased ATE's, but those cases must be analyzed separately from the normally assigned cases. In any case, researchers must be aware of the existence of specialized assignment *and* the court must have kept accurate records of such exceptions.

B. Differing Probability of Assignment

The Problem: As was discussed above, the basic model for computing an estimate of the ATE is the average outcome of interest of the treated subjects minus the average outcome of the non-treated subjects ($E[Y_i(1)|d_1 = 1] - E[Y_i(0)|d_1 = 0]$). This model assumes that the probability of being randomly assigned into the treatment group (often equal across groups, although it could be any value greater than 0 and less than 1) is constant for each of the subjects across the duration of the study. As a result, if the probability of being assigned into one or more of the treatment groups changes over the course of the study, this simple model generates a

[57] This process can be analogized to block random assignment, discussed in Chapter 4 of GERBER & GREEN, *supra* note 11.

potentially biased estimate for the ATE because the two groups that we are comparing, $Y_i(1)$ and $Y_i(0)$, are no longer identical in expectation.

An example using judicial assignment may be informative. Continuing on with our hypothetical study on the effect of being assigned to a male judge on the length of the defendant's criminal sentence, let us assume that the court system we are working in has 10 judges, 7 of whom are female and 3 of whom are male. Furthermore, let us assume that every new criminal case has a 10% chance of being assigned to each of the 10 judges. This means that any given case has a 30% chance of being assigned to a male judge and a 70% chance of being assigned to a female judge. If these probabilities remain constant over the course of the study, we can estimate the unbiased ATE of being assigned to a male judge simply by calculating the average sentence for all defendants assigned to male judges and subtracting the average sentence for all defendants assigned to female judges.

Now imagine that 2 of the male judges decide to escape the August heat by taking a vacation, so that during the month of August, there are only 8 judges—7 females and 1 male. With equal probability of assignment among the remaining 8 judges available to be assigned cases, the probability that any given criminal case is assigned to a male judge in August has now dropped from 30% to just 12.5%. Is this necessarily a problem? No, but only if we are willing to assume that the types of cases that come into the court during August are identical in every way to cases from the rest of the calendar year. If, however, August cases tend to be more complex, concern more serious crimes, or feature poorer defendants (any of which are likely related to our potential outcomes of interest), we might expect an increase (or decrease) in the expected sentencing lengths during that month, and because these August cases are more likely to be assigned to female judges, the resulting estimate of the ATE would be biased downwards.

i. Scheduling: The most common way in which the probability of judicial assignment is non-identical across time is in regard to scheduling. While most jurisdictions make a concerted

effort to keep each judge's docket running on schedule,[58] a number of issues can arise that may force the court to either lower the number of new cases being assigned to a particular judge (a decrease in the probability of being assigned) or take that judge off of the assignment schedule altogether (a probability of assignment of 0).[59]

While differences in assignment probability due to scheduling issues may seem like an innocuous occurrence, they become threats to unbiased causal inference when the change in probability is due to some factor related to potential outcomes. The vacationing judges discussed above is a potential example of this potential problem. If the cases assigned during time in which a judge is absent (and is not assigned cases) are different than cases assigned during the rest of the study period, an unbalanced estimate that includes that time period may be biased. Outside of vacations, scheduling can also alter assignment probabilities through sick-days, docket-backups, or changes in judicial responsibilities (e.g. a judge is assigned to take care of preliminary hearings for a week). Again, these alterations may not pose a threat to unbiased inference, but only if we assume that a judge's propensity for illness, inability to keep up with her docket, or likelihood of special assignments are unrelated to whatever potential outcomes we are concerned with.

ii. Consolidation of Codefendants: In many jurisdictions (at least 85% of the courts surveyed for this Chapter), codefendants' individual cases are consolidated to just one judge, generally to the judge assigned to the first defendant that comes up in the court system. Similar to scheduling effects, the impact of consolidating these cases can lead to changes in the probability of assignment, but only under specific circumstances.

---

[58] Not surprisingly, the most effective way of maintaining balance across judges is through random assignment itself. This scheduling utility and egalitarianism/fairness are the two primary reasons I was told that courts instituted random assignment procedures.

[59] This may result in the court reassigning some of a judge's already assigned cases to less-busy colleagues. This also poses a problem for causal inference, although it is different in nature than the differences in probability problem that I discuss in this Part. For a full discussion of this other problem, see *infra* Subpart III.C.i.

In courts systems that utilize a complete or rolling randomization, consolidation of a large number of codefendants can remove a judge from the assignment process for a substantial period of time. For example, the Harris County (Texas) Criminal Court randomly assigns judges to cases according to a type of complete random assignment where when a case comes in, their computer program identifies how many cases have been assigned to each of their criminal judges and the judge with the lowest number of assignments is assigned the case. If there are multiple judges with the lowest number of assignments, the computer randomly assigns the case to one of those judges with equal probability. In this system, each codefendant in a case counts as a single case, so a judge who is assigned a consolidated case with five codefendants, will jump up five spots in the assignment chart. Assuming the rest of the fourteen judges have a roughly equal number of assigned cases, this procedure means that the judge with the consolidated case will not be assigned any cases for five rounds, or 70 cases.

The Possible Solutions to Differences in Probability of Assignment: Much like the process of ensuring cases are randomly assigned, the first step in addressing differing probabilities of assignment is learning the probabilities themselves. Court representatives will, again, be the best source of information required to determine probability of assignment over the course of the time-period studied, but it may take substantial sleuthing to identify deviations from normal procedure. Researchers should ask whether caseloads are ever lightened when judge is behind on her docket, whether assignments continue to accumulate while judges are away for a sick-day or on holiday (and if so, whether there is a record of when specific judges were away from chambers) and should also attempt to reverse-engineer the exact probabilities at a given time using case-loads in the datasets. Alternatively (or additionally, if the researcher wants to verify the procedural descriptions of court personnel), researchers can "reverse-engineer" the actual probabilities of assignment for each case using the dataset itself. By comparing the number and type of cases that each judge has in a given period and across

periods, researchers can identify variations in assignment probability across individual judges and across time.

Once the probability of assignment for each case in the dataset is known, there is fortunately a simple solution to differences in probability of assignment that makes no sacrifices in regard to causal identification, even if such differences are associated with changes in potential outcomes. Through a technique called inverse probability weighting (IPW), the impact that each individual outcome of interest has on the estimated ATE can be re-weighted to reflect the likelihood of that unit being assigned to a particular judge or type of judge.

Using the above example of the vacationing judges, we can see how this is done in practice. We begin by dividing the dataset into time periods that have internally consistent probabilities across judges. These periods could be as short as a day (or shorter if probabilities of assignment shift mid-day), but to keep things simple, we will create two periods: the "normal" period (January to July and September to December), in which all 10 judges are in court and there is 30% chance of a given case being assigned to a male judge; and the "vacation" period, in which 2 of the male judges are on vacation, and there is only a 12.5% chance of being assigned a male judge. Instead of calculating the estimated ATE simply by subtracting $E[Y_i(0)]$ from $E[Y_i(1)]$, we re-weight the outcome of each case assigned to a male judge by 1/P (where "P" equals the probability of being assigned to a male judge for that case—.125 for cases assigned in August and .3 for cases assigned in any other month) and the outcome of each case assigned to a female judge by 1/(1-P) (1-P being the probability of being assigned to a female judge). We then derive the estimated ATE using the same approach as we did in Part II.[60] Notice that if the probability of being assigned to a particular judge or type of judge is zero at a given time, this would imply that cases assigned at that time will be re-weighted out of the calculation of the

[60] For a more comprehensive explanation of how IPW is done, see GERBER & GREEN, *supra* note 11, at 74-77.

estimated ATE (P would be equal to 1 or 0), which is the appropriate course given that assignment is non-random with such probabilities.

C. Post-Assignment Changes

The Problem: Generally, the judge assigned to manage a case remains the presiding authority over that case until it has concluded,[61] but complications and unforeseen circumstances can arise that lead to a change in the original judicial assignment. For example, the judge who has presided over the pre-trial portions of a case may be behind schedule and unable to run the trial itself, or a savvy attorney might seek a change in presiding authority after they see the initial judicial assignment. If we can assume that the propensity for post-assignment judicial changes are orthogonal to the potential outcomes of interest, these changes will not threaten the unbiased estimates of the ATE. However, under the many circumstances in which post-assignment changes are related to potential outcomes—using the male/female judges example, if guilty defendants are more likely than innocent defendants to seek reassignment from female judges—the resulting ATE estimate will be biased, even if the secondary assignment procedure is also done on a random basis.[62]

The experimental literature often refers to this sort of post-treatment change as crossover or non-compliance. The latter term is particularly helpful because it highlights the potential for post-assignment changes to reflect a sort of non-random self-selection out of the original treatment by one of the individuals involved with the case, generally the judge, the parties, or the attorneys. Building off of the notation used in previous parts, we can see why non-

[61] There are, of course, formal exceptions to this. Many courts have a special division or section that takes care of the arraignment portion of a case, and some, like Alameda County, assign out the multiple stages of a case (arraignment, pre-trial, trial) to different judges. Neither of these approaches prevent researchers from taking advantage of random assignment, although they make it more difficult to accurately model the entire assignment process and likely change the definition of the treatment.

[62] This result is not intuitive and has been missed by researchers in the past. While a random reassignment will ensure that the number and types of cases that a judge gains is orthogonal to pretreatment covariates, bias is nonetheless induced because the propensity for a judge to *lose* cases is likely a result of some important characteristic of the judge.

compliance might introduce bias. Remember that $Y_i(1)$ and $Y_i(0)$ represent the potential outcome when i is treated ($d_i = 1$) and untreated ($d_i = 0$), respectively. This basic notation assumes that assigned treatment and actual treatment are the same. Under a notation expanded to accommodate the possibility of non-compliance, however, the treatment *received* is potentially distinct from the treatment *assigned*, with the latter concept represented by z_i, where $z_i = 1$ when i is assigned to treatment, and $z_i = 0$ when i is assigned to control.

To combine the notation for treatment received and treatment assigned in a simple way (and because treatment received is ostensibly a function of treatment assignment), $d_i(z_i)$ will refer to a subject's treatment, given a particular assignment that the subject received. In a framework with only two treatment groups, there are four possible configurations of treatment and assignment: $d_i(z_i = 0) = 0$, or simply $d_i(0) = 0$, where a subject is assigned to control and is untreated; $d_i(0) = 1$, where a subject is assigned to control but is treated (one form of non-compliance); $d_i(1) = 1$, where a subject is assigned to treatment and is treated; and $d_i(1) = 0$, where a subject is assigned to treatment but is not treated (the second form of non-compliance).

Remember that because we are talking about *potential* outcomes, each subject is either compliant or non-compliant for each treatment group. A subject may, for example, comply with treatment if assigned to control ($d_i(0) = 0$) but be non-compliant if assigned to the treatment group ($d_i(1) = 0$). This, in turn, results in four possible "types" of subjects: *Compliers*,[63] who always comply with assignment ($d_i(1) = 1$ and $d_i(0) = 0$); *Never-Takers,* who are never treated, regardless of assignment ($d_i(1) = 0$ and $d_i(0) = 0$); *Always-Takers*, who are always treated

[63] The terminology used to refer to these four types of subjects varies, but the *Complier/Never-Taker/Always-Taker/Defier* labels are common and fairly intuitive. *See* Alexander Balke & Judea Pearl, *Nonparametric Bounds on Causal Effects from Partial Compliance Data*, Technical Report R-199, University of California, Los Angeles (1993); and Joshua D. Angrist, Guido W. Imbens, & Donald B. Rubin, *Identification and Estimation of Local Average Treatment Effects*, 62 ECONOMETRICA 467 (1996) (referring to the estimand as the LATE).

($d_i(1) = 1$ and $d_i(0) = 1$); and *Defiers*, who are treated when assigned to control and are untreated when assigned to treatment ($d_i(1) = 0$ and $d_i(0) = 1$).

In the court context, non-compliance occurs when a case ends up before a judge of a type other than the type of judge the case was originally assigned to.[64] In a study looking at the causal impact of having a female judge preside over a case, a case that was originally assigned to a male judge but was transferred to a female judge or assigned to a female judge but subsequently transferred to a male judge would be non-compliant, and a case that either retains the same judge throughout all proceedings or is transferred to a judge of the same type is referred to as compliant. If our dataset includes non-compliance in both the treatment and control groups, it suffers from two-sided non-compliance.[65]

Even with a dataset that features non-compliance, it may be tempting to estimate an ATE based off of the difference between those cases that ended up before female judges and those that ended up before male judges ($E[Y_i|d_i = 1] - E[Y_i|d_i = 0]$)—we are, after all, interested in identifying the impact of having a male or female judge on case outcomes. This would be unwise, however, because the distribution of subjects among male or female judges is likely no longer a function of random assignment. Expanded to incorporate non-compliance, the group of cases that were presided over by a male judge ($E[Y_i|d_i = 1]$) include both those cases that were presided over by a male judge and were initially assigned to a male judge ($d_i(1) = 1$) *and* those cases that were presided over by a male judge but were originally assigned to a female judge ($d_i(0) = 1$). The converse is true, of course, for the group of cases presided over by a female judge, which includes ($d_i(0) = 0$) and ($d_i(1) = 0$).

[64] Researchers may want to use a more restrictive definition of non-compliance where any deviation from the original assignment is non-compliance, even if the reassignment is to a judge of the same "type" (from one male judge to another, for example). As we see below, this would make the estimated treatment effects more difficult to calculate but may be necessary given specific research questions.

[65] Given the nature of the examples of post-assignment judicial changes outlined below, two-sided non-compliance is likely to exists in most in most court datasets.

In order to meet the requirement that the expected potential outcomes of cases eventually presided over by male judges equals the expected potential outcomes of cases presided over by female judges, the potential outcomes of the subjects who were not compliant with their random assignments to a female judge (likely a compilation of never-takers and defiers) must be the same as those of the subjects who were not compliant with their random assignment to a male judge (likely a compilation of always-takers and defiers).[66] As we will see from the examples of post-assignment changes below, this is unlikely.

i. Administrative Discretion: Often, the court administrators or the judge herself will identify some reason to make a change in judicial assignment after the initial assignment process has concluded. For example, if a judge feels that she is unfit to preside over a case, most commonly because of a conflict-of-interest, she may request to be removed from the case (recusal can also be prompted by party discretion). Alternatively, after a case has developed, a court may decide that it would be better handled by a different judge, possibly due to some attribute of the judge (bias or specialty), an attribute of the case (type or complexity), or a combination of both.

Administrative discretion can also be manifest in scheduling or experiential concerns. A court may choose to move some of the cases away from a judge who is having a hard time keeping up with her docket, or a chief judge might remove exceptionally complex or important cases from new and inexperienced judges. In either case, because the decision to make a post-assignment change is based in part on the characteristics of the judge or the case, such decisions are likely to create uneven sets of non-compliers, resulting in inferential bias.

ii. Party Discretion: While the original assignment process is always under the discretion of the court system, rules and procedures sometimes give opportunities for the parties

[66] By implication, the subjects who were compliant with their random assignment to a female judge must also have the same expected potential outcomes as those who were compliant with assignments to male judges.

involved in the case to ask for post-assignment changes in judicial oversight. In some states, parties to a case can request a judicial transfer via a peremptory challenge, in which a party asks for a new judge, often without needing to cite a particular reason for the request.[67] Most states allowing peremptory challenges restrict each party to one such request over the life of the case, although a few allow multiple challenges for specific types of cases.[68] Similarly, attorneys may make substantive pleas, which are normally used to move a case to a judge with a particular expertise but may strategically be used when the original judge is perceived to be unfavorable to a client's case.

Additionally, every state allows parties to request a judicial recusal if the judge is unable to adjudicate the case in an unbiased manner. Recusal, as was noted above, is often up to the judge's or administration's discretion, but these decisions may ultimately be impacted by the recusal requests of attorneys or third parties.

More nefariously, attorneys may seek to circumvent random and as-if random assignment processes by strategically submitting cases in a manner or at a time that will increase the probability that a favorable judge will be assigned to the case. In a particularly interesting example of this, a prestigious Chicago law firm was caught judge-shopping by filing multiple, identical lawsuits and hoping that the random assignment process used by the Chancery Court would yield at least one friendly judicial assignment. They would then drop the remaining cases assigned to less-desirable judges. [69]

The Possible Solutions to Post-Assignment Changes: There are two primary approaches to addressing non-compliance due to post-assignment judicial changes, both of

[67] For a general discussion of peremptory challenge (and recusal) procedure and constitutionality, see Chapter 26 in RICHARD FLAMM, JUDICIAL DISQUALIFICATION: RECUSAL AND DISQUALIFICATION OF JUDGES (2d ed. 2006); and Nancy J. King, *Bastion for the Bench – Regulating the Peremptory Challenges of Judges*, 73 CHI-KENT L. REV. 509 (1998).

[68] For a fairly recent examination of which states feature these judicial peremptory challenges, see Chapter 27 in *id.*

[69] Andrew Fegelman, *Law Firm is Fined, Scolded*, CHI. TRIB., April 5, 1994.

which come with methodological costs. Under the first option, the researcher may essentially choose to define away non-compliance by changing the causal effect of interest from the ATE of a case being presided over by a male judge ($E[Y_i|d_i = 1] - E[Y_i|d_i = 0]$) to the ATE of the case being *assigned* to a male judge ($E[Y_i|d_i(1)] - E[Y_i|d_i(0)]$)—a treatment effect often called the intent to treat effect, or ITT. The ITT "avoids" non-compliance because it is concerned only with the original case assignment, which cannot be changed, regardless of how many times the case is re-assigned. The downside to using the ITT is, of course, that the treatment effect of assignment itself is rarely of primary interest to researchers. However, the extent to which an estimated ITT differs from an estimated ATE depends on the frequency of non-compliance due to post-assignment changes, so circumstances in which such changes are ostensibly rare and will consequentially produce minimally divergent ITT estimates.[70]

The alternative to the ITT, the complier average causal effect (CACE),[71] does provide researchers with an estimate of the treatment effect, albeit only for cases that are compliers ($E[\{Y_i(d_i = 1) - Y_i(d_i = 0)\}|d_i(1) = 1]$) and only under a number of important assumptions. While this approach makes analytic sense in a number of research contexts, it is likely to provide only moderate value for researchers studying the impact of case assignment, as such studies are generally concerned with outcomes for all types of cases or subjects, not just those who do not move between judges. For this reason (and because the process of estimating the CACE is fairly involved), this procedure is not explained in detail here. In short, it requires the researcher to assume that none of the cases in the dataset are "defiers" ($z_i \neq d_i$, regardless of initial assignment—an assumption also called monotonicity), after which the ITT is divided by the

[70] Non-compliance can presumably be identified by comparing original assignments to the assignments at the conclusion of the case, although such information may be available only through specific court records as opposed to the general docket data.

[71] The complier average causal effect is also referred to as the local average treatment effect, or LATE. *See* Angrist, Imbens, & Rubin, *supra* note 63 (referring to the estimand as the LATE).

average $d_i = 1$ minus the average $d_i = 0$. Researchers hoping to address non-compliance using the ITT or the CACE should refer to the number of particularly helpful guides on the topic.[72]

D. Missing Outcomes (Attrition)

The Problem: The most statistically troubling problem that researchers will likely encounter when utilizing random judicial assignment is missing outcomes. Also referred to as attrition (the term used in this Chapter), missing outcome data can have devastating consequences for unbiased causal analysis when missingness is correlated with potential outcomes. Making matters worse, missing outcome data in the court context is unavoidably common, meaning that it is essential for researchers to understand the implications of attrition and be able to identify when and to what extent it exists in a given dataset.

Remember that the inferential power of random judicial assignment comes from its ability to create sets of cases that are statistically equivalent in expectation, thus allowing us to compare the outcomes of those cases as a function of the treatment. Or, using the earlier notation:

$$E[Y_i(1)|d_1 = 1] = E[Y_i(1)|d_1 = 0],$$[73]

meaning the expected treated outcome for cases in the treatment group is equal to the expected treated outcome for cases in the control group (the same would be true for expected non-treated outcomes in both groups).

When considering the role that attrition plays in our dataset, it can be helpful to think of missingness as a potential outcome itself. Just like $Y_i(1)$ and $Y_i(0)$ represent the potential outcomes of our outcome of interest given assignment to treatment and control, $r_i(1)$ and $r_i(0)$ represent whether a given data point or case will be missing or not given assignment to

[72] For a more comprehensive discussion on deriving the ITT and CACE, see Chapters 5 and 6 in GERBER & GREEN, *supra* note 11; and Chapters 23 and 24 in in IMBENS & RUBIN, *supra* note 12.

[73] In the interest of simplicity, this equation assumes that assignment (z_i) always matches treatment (d_i).

treatment or control, where $r_i(1 \text{ or } 0) = 0$ means that the outcome of interest (Y_i) will be missing and $r_i(1 \text{ or } 0) = 1$ means it will be non-missing, or observed.

Going back to the equation above, it then follows that the expected outcome of interest for subjects assigned to the treatment group is a combination of those subjects who have missing outcomes under treatment and those who have observable outcomes under treatment:

$$E[Y_i(1)] = E[r_i(1)] * E[Y_i(1)|r_i(1) = 1] + \{1 - E[r_i(1)]\} * E[Y_i(1)|r_i(1) = 0],$$

where the $E[r_i(1)]$ and $\{1 - E[r_i(1)]\}$ act as switching mechanisms since only one outcome (missing or not) is observed. The same can be done for the control groups, $E[Y_i(0)]$, and replacing both groups with the expanded notation results in

$$E[r_i(1)] * E[Y_i(1)|r_i(1) = 1] + \{1 - E[r_i(1)]\} * E[Y_i(1)|r_i(1) = 0] =$$

$$E[r_i(0)] * E[Y_i(0)|r_i(0) = 1] + \{1 - E[r_i(0)]\} * E[Y_i(0)|r_i(0) = 0].$$

In a dataset or study with no attrition, this equation would simplify back down to our original ATE equation because the potential for missing data on both sides would be 0. However, if attrition does exist (and as we will see, it almost always does in the context of court data) we are required to make the strong assumption that the expected value of the missing data on each side of the equation is exactly the same, which is difficult to do because such outcomes are, by definition, not observable.

To illustrate this using a slight variation of our ongoing example, imagine that we are looking to identify the impact of being assigned to a male judge on the outcomes of divorce cases but have noticed that many cases that started in court are resolved through private dispute settlements, the outcomes of which we do not have access to. Attrition clearly exists here, but because the outcomes of the missing cases are, by definition, unobserved, we are left with only the observed cases and are obliged to assume that the missing outcomes of cases that are assigned to male judges are the same in expectation to the observed outcomes of cases assigned to male judges, or that:

$$E[Y_i(1)|r_i(1) = 0] = E[Y_i(1)|r_i(1) = 1].$$

The same must also be true, of course, for observed and missing cases assigned to female judges.

These assumptions might be true. It is possible that the propensity for a case to be resolved through settlement is completely independent of potential outcomes of that case. But suppose that outcomes are correlated with the propensity for attrition and that attrition is more common for cases assigned to female judges (possibly because attorneys in our sample feel that male judges are relatively more favorable towards their clients than female judges). This would mean that the set of observable cases before male judges would have more (and by extension, different) cases than the set of observable cases before female judges, and because the missing cases are unobservable, we cannot know whether the resulting estimate of the ATE is the product of the gender of the judge the cases were assigned to or the difference in case composition.

As we will see below, settlements are not the only source of attrition-based bias. Alternative forms of strategic lawyering such as plea-bargaining and re-filing as well as administrative error, selective publication of case outcomes, and data corruption are all potential hazards for researchers hoping to utilize random judicial assignment.[74]

i. Plea Bargains, Settlements, and Strategic Behavior: A large number of criminal cases never make it to a judicial or jury determination because the defendants choose to plead guilty for a reduced punishment (90-95% by some measures).[75] Likewise, an increasing proportion of civil cases are resolved through non-judicial means, despite starting out as judicially managed cases in the courts.[76] While plea bargains and settlements do not inherently lead to attrition

[74] It may be more intuitive to understand the potential bias introduced by attrition in the medical trials, where outcomes of interest generally relate to health and attrition takes the form of death. For a short but elucidating discussion of this, see Samuel L. Brilleman, Nancy A. Pachana, & Anette J. Dobson, *The Impact of Attrition on the Representativeness of Cohort Studies of Older People*, 10 BMC MED. RES. METHODOLOGY 71 (2010).

[75] *See* Lindsey Devers, *Plea and Charge Bargaining*, Bureau of Justice Assistance: US Department of Justice, Order No. 2008_F08151 (2011).

[76] Settlement rates seem to vary quite widely depending on the case type. Eisenberg and Lanvers found that in two Federal District courts (Eastern District of Pennsylvania and Northern District of Georgia) the rates were as low as 27.3% for constitutional tort cases and as high as 87.2% for regular tort cases. *See* Theodore

(some outcomes of interest are available regardless), many of the questions that scholars are interested in answering rely on outcomes that are dependent on case data or court records which are not available for cases that are resolved out of the courtroom. In these cases, estimated treatment effects will be biased if the attrition is related to the potential outcomes in any way.

For example, sophisticated lawyers may have certain perceptions (accurate or not) regarding how punitive or difficult certain judges are and consequentially advise their clients to take a plea deal more often under some judges than under others. Over time, the "tough" judges would have a set of measurable outcomes that are made up of relatively fewer defendants with sophisticated lawyers than the "easy" judges, and if we believe that sophisticated lawyering is related to outcomes, then this selection effect will induce significant bias in a study's results. In addition to avoiding a final determination by the courts through plea bargains or settlements, attrition might also result from civil plaintiffs strategically dropping their cases altogether.

When attrition is the result of strategic behavior on the part of parties or attorneys, researchers might consider using attrition itself as the outcome of interest. Whether or not a criminal defendant pleads guilty, for example, is an interesting outcome to measure, and given the effect that judicial characteristics such as gender and race seem to have on more traditional outcomes such as sentencing, we might imagine that those same characteristics impact the propensity for plea bargaining. This outcome may not be exactly what the researcher wished to measure, but it will be valuable as an unbiased measure of the treatment effect.

ii. Selective Publication of Data, Decisions, and Opinions: Most court adjudicated case outcomes are legally required to be publicly available.[77] However, certain case data are

Eisenberg & Charlotte Lanvers, *What is the Settlement Rate, and Why Do We Care*, 6 J. of Emp. Legal Stud. 111 (2009).

[77] Each state has different rules regarding public records, but all of them make the majority of cases available. For a full list of public record laws, see FOIA Advocates, "State Public Record Laws," available at http://www.foiadvocates.com/records.html.

selectively kept off record. A court may have privacy concerns about cases that involve minors or cases dealing with classified information and choose not to make dockets, documents, procedural decisions, or—less frequently—case outcomes publicly available. Similarly, some courts (generally courts of appeals) write case opinions that outline the reasoning behind the judicial disposition, which are used by researchers to measure legal reasoning or policy influence. Because judges generally have full dockets, they do not write—or at least do not publish—opinions for many of the cases assigned to them, presumably spending their time on cases that deal with novel questions or have particularly high precedential value. Some have even suggested that judges take advantage of publication to push certain political or legal ideologies, in which case, attrition due to non-publication is highly likely to relate to potential outcomes.[78]

The Possible Solutions to Missing Data (Attrition): Before addressing attrition, the researcher must be able to identify whether and to what extent her dataset features missing data. This process can be difficult because court data, particularly state court data, is often unreliable, and even previously analyzed datasets may have undocumented attrition. As a first step, researchers should talk with a court's data-specialist about whether a given dataset includes all cases assigned during that time period.[79] Assuming that courts have a systematic method for assigning case numbers (something that can be confirmed with the court representative), missing cases should be identifiable based on numerical gaps in an ordered list of those case numbers. The researcher might also run a balance check on a number of pre-treatment variables that, under random assignment, should be equal across treatment groups. Caution should be used when interpreting these results, however, because imbalances might be

[78] *See* David S. Law, *Strategic Judicial Lawmaking: Ideology, Publication, and Asylum Law in the Ninth Circuit*, 73 U. Cin. L. Rev. 817 (2012).

[79] It is key to be specific with these questions. *See infra* Subpart IV.C for thoughts on how to approach these discussions.

due to other de-randomizing effects such as post-assignment judicial changes (see Subpart III.E for more on balance checks).

There is no simple way to deal with attrition once it is identified. Under ideal circumstances, the attrition is not related in any way to the outcomes of interest, resulting in a condition sometimes referred to as *missing independent of potential outcomes* (or MIPO). While MIPO cannot be empirically verified,[80] certain procedural events—particularly those that are non-deliberate or haphazard (lost or missing data are common examples)—are ostensibly less likely to feature a connection between the propensity for attrition and the potential outcomes of interest, allowing for more reasonable assumption-making. Similarly, the attrition might be independent of potential outcomes given a known pre-treatment covariate (appropriately called MIPO|*X*), in which case, the propensity for missingness is unrelated to potential outcomes for a subgroup of the subjects. If researchers have good reason to believe this is true (again, it cannot be empirically verified), they can use the average observed outcomes of interest in each subgroup as a proxy for the missing data (making sure to re-weight the data using IPW), resulting in an unbiased estimate of the ATE.[81]

More often, however, attrition cannot be assumed to be independent of outcomes of interest. Under these circumstances, there is a limited set of approaches that the researcher can take. She may, of course, choose to just estimate the treatment effect despite missing outcomes, hoping that the bias resulting from attrition is not significant, but this will ultimately render the probative value of the results largely undeterminable. More appropriately, the researcher can fill

[80] Because identifying the source of attrition is a causal question, researchers hoping to verify MIPO or MIPO|X will run into the same inferential problems as we discussed in Subpart II.A, *supra*. However, researchers willing to make a number of assumptions can test for obvious signs of attrition by comparing rates of attrition between the treatment groups (difference in rates suggest that attrition may be a function of assignment) or calculating the rates of attrition within a series of pre-treatment covariates. *See* Subpart III.E for a discussion on balance tests.

[81] For a more detailed discussion of MIPO and MIPO|X, see GERBER & GREEN, *supra* note 11, at 219.

in the missing outcomes with estimated values, either by drawing bounds around such estimates or, preferably, by ascertaining accurate estimates with additional data collection.

Filling in missing outcomes with educated estimates, may seem to be an unsuitable methodological approach to dealing with attrition, but a number of estimation tactics require only minimal assumptions. The most conservative approach to estimating missing data is to use extreme value bounds, in which a researcher assigns each missing outcome to the lowest and highest possible outcome values. For example, if the outcome of interest is length of sentence, each missing outcome will be assigned both the minimum (presumably 0) and maximum (some value above 0) possible sentence lengths for a given crime, creating a set of bounds within which all possible estimated ATEs will exist. If the outcome of interest is binary, such as guilty or not guilty, all missing values will be assigned both a 0 and a 1. A more moderate approach to bounded outcomes would be to assign missing outcome values to the lowest and highest values *observed* in the non-missing set, and a more liberal approach (one that relies on an assumption of monotonicity) would be to use a trimming technique on the observed data.[82] How conservative the researcher wants to be regarding these bounds will likely have a substantial impact on the estimated standard errors of the estimated ATE and the size of the bounds themselves, often increasing them enough to preclude any statistically significant findings.

If additional resources are available, researchers should consider expending an extra effort in gathering a random sample of the missing outcomes. Commonly utilized in survey research as a solution to non-responses, double sampling can provide researchers with a representative sub-sample of outcome measures that can then be applied to the missing data as a whole. While this approach is likely more appropriate for researcher-controlled randomized experiments, in which the temporal and analytical distance between treatment and empirical

[82] *See* Peter M. Aronow, Donald P. Green, & Donald K. K. Lee, *Sharp Bounds on the Variance in Randomized Experiments*, 42 ANNALS STAT. 850 (2014); and David Lee, *Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects,* 76 REV. ECON. STUD. 1071 (2005).

assessment is small, it should be considered by researchers utilizing random judicial assignment as well and may be combined with other methods such as bounded outcomes.[83]

E. Verifying the Model of Assignment Procedures

Once a researcher is confident that she has identified all of the various de-randomizing events that might be featured in a particular dataset, she should verify the randomness in her model using a mixture of statistical tests and qualitative indicators. As I explain in Part IV, it became clear while surveying administrators and judges that many courts are hesitant to disclose informal deviations from the regular assignment process for fear of perceptions of impropriety,[84] even when evidence of non-randomness in purportedly randomized systems exists.[85] Alternatively, the administrators or the researchers may have simply overlooked certain procedures during the initial data collection. As a result, researchers interested in utilizing random judicial assignment should not rely solely on the list of procedures provided by the court system.

In this Subpart, I briefly outline two approaches that researcher can take to verify the assignment procedures featured in their court data: statistical balance tests and informal "tells" of non-random deviation.

i. Formal Statistical Tests: As we discussed above, random assignment produces comparison groups that are equal in expectation, or:

$$E[Y_i(1)|d_i = 1] = E[Y_i(1)|d_i = 0] = E[Y_i(1)] \text{ and}$$

$$E[Y_i(0)|d_i = 1] = E[Y_i(0)|d_i = 0] = E[Y_i(0)].$$

[83] *See* Peter M. Aronow, Alexander Coppock, Alan Gerber, Donald P. Green, & Holger L. Kern, *Combining Double Sampling and Bounds to Address Non-Ignorable Missing Outcomes in Randomized Experiments*, 25 POL. ANALYSIS 188 (2015).

[84] *See* press articles in *supra* note 5 for examples of the public's concern with seemingly non-random judicial assignment.

[85] Chilton and Levy provide evidence of non-randomness in a number of U.S. Circuit Courts, even against the insistence of those courts that the process is random (Chilton & Levy, *supra* note 5; and Levy, *supra* note 5).

This expected equality, or balance, between treatment groups extends to all pre-treatment covariates, even those that are unrecognized by researchers or altogether unobservable or undocumented in the dataset. Statistical balance tests (also called randomization tests) can provide researchers with evidence that a particular court dataset features randomly assigned cases by comparing the covariate distributions in the various treatment groups. Under the assumption that these two groups are equivalent, their covariate distributions should be statistically indistinguishable.

These types of tests generally take one of two forms in the literature: a t-test or a bootstrapping method. A t-test (more formally known as a Student's t-test) uses the treatment groups' means, standard deviations, and sizes to measure the overlap in the distributions of a particular covariate (presumably one that is prognostic of the outcome of interest). Statistically "balanced" distributions will overlap to a substantial degree, while non-overlapping distributions may be indicative of a non-random process. These tests, while simple, are limited by the distributional assumptions built into them—most importantly that the underlying distributions are normal. Bootstrapping methods, on the other hand, do not make any such distributional assumptions, relying instead on actual data produced by repeated simulations of the case assignment protocols. A number of recent papers showcase the value of such an approach while simultaneously highlighting the modeling and computational difficulties inherent in them.[86]

Running these statistical tests come with three important caveats. First, they are truly prognostic only inasmuch as the assignment models that they are based on are correct—an incorrectly modeled assignment protocol could produce test results that indicate non-randomness when the actual cases were randomly assigned or, conversely, produce results that are supportive of random assignment but miss some important source of non-random bias.

[86] *See, e.g., id.*

Second, both methods described above rely on the assumption that the distributions of the pre-treatment covariates tested are reflective of all the possible sources of bias. Bias-inducing attrition caused by strategic attorney behavior, for example, may not be related to the gender of the defendant and would be missed by a balance test that looks only at the gender distribution in the treatment groups. Third, an imbalance in the distribution of a covariate is not necessarily an indication of non-randomness. Random assignment only ensures that the treatment groups will be equal *in expectation*, so imbalances may merely be the result of random chance, particularly if the researcher is using a small sample size or testing a large number of covariates.[87] Statistically significant imbalances on a large number of prognostic covariates constitutes stronger evidence of non-randomness, which should be addressed by digging deeper into the details of the assignment process.[88]

In any case, just as it would be unwise for a researcher to rely entirely on the assignment protocols outlined by court administrators or earlier studies, attempting to identify the various de-randomizing events described above using just a balance test would potentially miss the nuanced—and occasionally significant—sources of statistical bias that are featured in nearly all court systems and, in some cases, may prompt researchers to prematurely abandon an empirical pursuit because of seemingly non-random data. Instead, these tests should be understood as supplements to the sort of *pre*-considerations suggested in this Chapter.

---

[87] Researchers should remember to account for the multiple comparison problem by appropriately adjusting the resulting standard errors. For a discussion of this problem and how it can be addressed, see M. Aickin & H. Gensler, *Adjusting for Multiple Testing when Reporting Research Results: The Bonferroni vs Holm Methods*, 86 AM. J. PUB. HEALTH 726 (1996). Alternatively, researchers may implement procedures that combine multiple balance tests into one single procedure. *See* Ben B. Hansen & Jake Bowers, *Covariate Balance in Simple, Stratified, and Clustered Comparative Studies*, 23 STAT. SCI. 219 (2008).

[88] Many suggest that the results of balance tests can also be used to re-calibrate the data on the offending covariates (e.g. through post-stratification), although the efficacy of such an approach is questionable. *See* Diana Mutz & Robin Pemantle, *The Perils of Randomization Checks in the Analysis of Experiments*, University of Pennsylvania Working Paper (2011); and Stephen Senn, *Testing for Baseline Balance in Randomized Clinical Trials*, 13 STAT. MED. 1715 (1994).

ii. Informal Indicators of Non-Randomness: In addition to the formal statistical tests described above, researchers might also consider a more qualitative approach to verifying that the assignment procedures described to them by court administrators are truly reflective of the day-to-day workings of the court. For example, an initial assignment procedure that divvies up cases based on simple randomization (each case has an X% probability of going to each of the judges) should, somewhat counter intuitively,[89] result in a number of "runs," or consecutive cases being assigned to the same judge. The likelihood of these runs depends of course, on the length of the run, the number of available judges, and the probability of being assigned to each of those judges, but one would expect, for example to see five, six, or even seven cases consecutively assigned to the same judge in large trial court datasets. Conversely, if the assignment protocol purportedly follows a complete- or "wheel"-based system, in which cases are assigned based off of a recurring list of judges or assigned to judges one-by-one until each judge has a case, one would not expect to see any runs at all (or at least no runs greater than two). Other seemingly non-random events such as all-minority judicial panels, long periods of non-assignment to some judges, and temporary imbalances between judges on salient pre-treatment covariates can also serve as "tells" for truly random assignment procedures.

F. "De-Randomizing" Events Checklist

While the above recommendations do not represent a comprehensive collection of all the considerations that should be made when using random judicial assignment to make causal claims, they focus the most common events that may introduce non-random bias into an otherwise workable dataset. Assuming that the researcher has verified that the general

[89] When one understands how random assignment works, of course, these "runs" are not counterintuitive at all, but for many (possibly including the court staff who makes the assignments), they may seem indicative of a problem in the assignment process and be cause for administrative discretion. The "Hot-hand Fallacy," named after the traditional belief among basketball fans that making a basket increases the likelihood that the next shot will go in, provides an interesting example of how random consecutive runs can seem legitimately non-random. *See* Thomas Giovich, Robert Valonne, & Amos Tversky, *The Hot Hand in Basketball: On the Misperception of Random Sequences*, 17 COGNITIVE PSY. 295 (1985).

assignment process is random, the following checklist summarizes the basic steps that researchers and readers engaging in such an enterprise can take:

Figure 1: De-randomizing Events Checklist

A. Ensure That General Assignment Process is Random or As-if Random
- ☐ *Balance Check:* After conducting a randomization check (t-test or Fisher's Exact Test, see Subpart III.E), do treatment groups appear balanced on pre-treatment variables?
- ☐ *Non-random "Tells":* Probabilistically, random assignment should naturally produce certain assignment patterns, such as runs of cases successively assigned to the same judge or all-minority panels. The absence of these "tells" may be an indication of non-random discretion by court administrators.

B. Identify Deviations from Initial Process: Are there any institutionalized, non-random exceptions to the generally random assignment procedures?
- ☐ *Specialization*: Do some judges specialize in certain case types?
- ☐ *Administrative Discretion:* Do the chief judges or administrators ever have discretion in the assignment of individual cases? (e.g. because of case complexity or length)
- ☐ *Party's Past Participation:* Are case assignments based on a party's past participation in criminal or civil trials? (e.g. repeat offenders, pending cases, or case extensions)
- ➢ **Possible Solutions:**
  - Ensure that the initial assignment process is accurately mapped.
  - Exclude cases that were assigned using non-random procedures from analysis.

C. Account for Changes in Probability of Assignment: Are there any variations in the probability of assignment over the period that the study is concerned with?
- ☐ Is the probability of assignment not recorded by the court system?
- ☐ Are some cases assigned in groups ("block" randomized)?
- ☐ *Scheduling:* Are caseloads adjusted based on a judge's schedule? (e.g. vacations)
- ☐ *Party Consolidation:* Does party consolidation impact probability of future assignments?
- ➢ **Possible Solutions:**
  - Use the assignments in the dataset to "reverse-engineer" (see Subpart III.B) estimated probability of assignment (but be wary of post-assignment changes and attrition).
  - Implement IPW (inverse probability weights) in statistical tests.

D. Account for Non-compliance: Are there any post-assignment changes to the initial assignment?
- ☐ *Administrative Discretion:* Does the court or the judge ever change the assignment?
- ☐ *Party Discretion:* Does the party have any influence on re-assignment? (e.g. preemptory challenge)
- ➢ **Possible Solutions:**
  - Ensure that the court records include original assignments, before deviations occur.
  - If non-compliance is minimal, estimate the ITT (intent to treat effect)
  - If assuming monotonicity is plausible, estimate the CACE (complier average causal effect).

E. Account for Attrition: Does the dataset have missing outcomes or missing data?
- ☐ Is the missing data an outcome of interest?
- ☐ *Strategic Resolutions:* How often are cases resolved out of court? (e.g. plea or settlement)
- ☐ *Selective Publication:* Does the court or judge have discretion regarding the publication of data? (e.g. selective writing/publication of opinions)
- ➢ **Possible Solutions:**
  - Check whether attrition is independent of potential outcomes by testing on prognostic pre-treatment variables.
  - Fill in missing outcomes and draw bounds around point estimates.
  - Gather missing data through random double sampling.
  - Treat missingness as an outcome of interest.

F. Re-Conduct Balance Checks While Accounting for Checked Events: After mapping initial assignment process and accounting for "de-randomizing" events, do cases appear to be randomly assigned based on empirical balance tests?

IV. A Survey of Assignment Procedures Used in U.S. Criminal Courts

To supplement the methodological assertions outlined above, I conducted a comprehensive survey of the judicial assignment procedures of the 30 largest state criminal courts in the United States. The results of this survey are valuable in at least three distinct ways. First, the questions court officials were asked correlate with most of the "de-randomizing" elements discussed in Part III, and the resulting data demonstrates the extent to which those events exist and should be taken seriously by researchers. Second, at the time this Chapter was written, there was no database of state court procedures comprehensive enough to address the needs of scholars hoping to utilize random assignment of cases. This data, therefore, should serve as a practical starting point for researchers looking for venues with assignment procedures amenable to unbiased causal inference.[90] Finally, the survey results should be useful to readers and reviewers who want to confirm the existence or non-existence of certain assignment procedures in courts that are featured in other empirical analyses that purport to utilize random assignment.

In this Part, I describe the sample of courts featured in the survey and outline the survey methods and questions. I then describe the results of the survey, apply them to the analysis in Part III, and share some helpful insights I gained while collecting the data.

[90] While this data constitutes the most comprehensive and accurate collection of state-level judicial assignment procedures that I am aware of, all the information was collected via phone conversations and emails and, as a result, should not be relied on to asses court data. In addition, this data only reflects assignment procedure at the time of the survey (May 2014 to May 2015), not past procedure—knowledge of which is necessary when using data from that period—or the procedure used at the time this is read—which may be different than the procedure reported in this Chapter. Researchers interested in analyzing data in any of these courts should confirm the data in this Chapter themselves. Not only will this ensure that any causal claims are empirically sound, personally learning about the assignment procedures is an excellent way to get to know the court structure and staff, an essential component in the data collection and analysis stages of these research endeavors.

A. The Sample

The survey sample presented below includes the criminal courts of the 30 most populous U.S. counties.[91] Criminal court size was determined using county-level population statistics[92] based on the U.S. Census Bureau's 2013 estimates.[93] Not surprisingly, the majority of these courts are located in the larger states—California (8 counties), New York (6), Texas (4), and Florida (4)—or larger cities—Mesa/Phoenix, Boston, Detroit, Las Vegas, Cleveland, Philadelphia, and Seattle. The average population of the counties is 2,451,488, with the largest being Los Angeles County (population 10,017,068; 431 judges[94]) and the smallest being Cuyahoga County, OH (population 1,263,154; 34 judges[95]).

The data resulting from this sample will likely be the most valuable for researchers interested in utilizing random judge assignment in their studies because the largest courts have the highest number of judges (increasing the likelihood of observing diversity in judicial characteristics and behavior), and the most cases (data points). As a result of their size, however, these court systems also tend to be more administratively complex, increasing the number of problematic procedures that were outlined above in Part III.

---

[91] The list of counties in the large sample are: Los Angeles, CA; Maricopa, AZ; Orange, CA; Riverside, CA; San Bernardino, CA; Sacramento, CA; Miami-Dade, FL; Broward, FL; Middlesex, MA; Wayne, MI; Clark, NV; Kings, NY; New York, NY; Suffolk, NY; Nassau, NY; Cuyahoga, OH; Philadelphia, PA; Tarrant, TX; King, WA; Bexar, TX; San Diego, CA; Harris, TX; Santa Clara, CA; Hillsborough, FL; Bronx, NY; Cook, IL; Alameda, CA; Palm Beach, FL; Queens, NY; and Dallas, TX.

[92] Because the vast majority of state trial court jurisdictions are defined according to county lines, I used county-level population as a rough measurement for criminal court "size." Defining court size using geographical area, number of judges, or number of cases would also have been equally plausible alternatives.

[93] U.S. Census Bureau: Population Division, March 2014, "Annual Estimates of the Resident Population for Counties: April 1, 2010 to July 1, 2013," *2013 Population Estimates*, retrieved April 17, 2014 (http://www.census.gov/popest/data/counties/totals/2013/CO-EST2013-01.html).

[94] Based on numbers reported on the official California Courts website, retrieved April 17, 2014 (http://www.courts.ca.gov/2948.htm).

[95] Based on numbers reported on the official Cuyahoga County website, retrieved April 17, 2014 (http://cp.cuyahogacounty.us/internet/Judges.aspx).

B. Survey Method and Questions

All surveyed counties were contacted from May 2014 to April 2015. Three counties—Palm Beach County, Los Angeles County, and Queens County— either declined to participate in the survey or were unable to find a representative who could answer the survey questions.[96] While all counties were originally contacted over the phone, some answered the survey questions through email or post. The individuals who answered the questions were generally the supervising/chief judge, the lead clerk, or the department coordinator, although a number of courts using complex computer algorithms to make assignments referred me to their information technology specialist. In order to ensure accurate information on the courts' assignment procedures, I always asked to speak with the individual who "directly handles the case assignment procedure."

The survey itself was a semi-structured questionnaire focused on three main inquiries.[97] First, I determined whether the initial assignment process was random. While this may seem to be a simple question, the exact method of assignment turned out to be quite difficult to establish. Most court representatives, while eager to help, were oblivious to the technical details of the assignment process, and even those specifically tasked with this responsibility were often not fully aware of how exactly cases were divided amongst the judges (referring to "the computer" as an explanation was common). Second, I asked the subjects about any additional events that occur during and after judicial assignment, including most of the de-randomizing elements discussed above. Because the questions associated with these events were relatively straightforward (e.g. "are repeat offenders assigned to their previous judge?"), subjects generally knew the necessary details. Finally, I asked how thoroughly the probability of assignment and post-assignment procedures were recorded.

---

[96] This is generally because none of the staff with public phone numbers nor the individuals they referred us to knew who was in charge of case assignment.

[97] While all subjects were given the same general questions, the exact wording and order changed depending on whom the subject was and how familiar they were with the technicalities of the case assignment process.

C. Survey Results

Below I present and discuss the aggregate data from the court survey, focusing specifically on the extent to which the events discussed above play out in the largest criminal courts. I find that two-thirds of the surveyed courts utilize random or as-if random initial assignment, but also that nearly all of the "de-randomizing" events discussed in Part III are commonplace. I also share some additional, non-empirical insights gained through the survey process, including suggestions for researchers who wish to verify this survey data themselves or classify the assignment protocols in additional courts. The tables in this Part contain only basic data (generally binary indicators) on the judicial assignment protocols.

i. Types of Assignment: The data from the survey show that the distribution among assignment types in the largest criminal courts is roughly equal, with 11, 7, and 9 courts utilizing random, as-if random, and non-random procedures, respectively (see Tables 1a-1c, below). Of course, the distinction between these assignment procedures depends on the assumptions that the researcher is willing to make and the outcomes of interest she is studying, but for organizational purposes, I have categorized any procedures that assign based on the schedule of the judge or the geographic area in which the crime was committed as non-random and procedures that assign according to a reoccurring, non-random list (often called a "wheel") as as-if random. Among the 11 courts that use random assignment, all but one (Nassau County, which physically pulls names from an envelope) performs the random assignment using computer-generated numbers. These numbers are manifest either as stand-alone identifiers or are incorporated into the official case number itself.

Every one of the courts that employ either random or as-if random assignment features at least one of the non-random procedures discussed in Subpart III.A. All of the courts that use random assignment have some judges who specialize in certain types of cases, as do all but two of the courts that use as-if random assignment. While assignment based on case type or geography are normally dispositive of truly random assignment, some courts randomly assign

Table 1a: Courts with Random Assignment Protocols

Random Assignment: Cases are allocated independently of any value, characteristic, or variable other than an exogenous assignment mechanism.				
County Name (State)	**Assignment Mechanisms***	**Specialization (Type)****	**Repeat Offenders*****	**Pending Case/ Probation*****
Clark (NV)	Computer and Geographic Area	Yes (Multiple)	No Protocol	Assigned to Previous Judge
Cool (IL)	Computer	Yes (Multiple)	No Protocol	No Protocol
Cuyahoga (OH)	Computer	Yes (Multiple)	NA	No Protocol
Dallas (TX)	Computer	Yes (Multiple)	No Protocol	No Protocol
Harris (TX)	Computer and List	Yes (Multiple)	No Protocol	No Protocol
Hillsborough (FL)	NA	Yes (Multiple)	No Protocol	No Protocol
Maricopa (AZ)	Computer	Yes (Multiple)	No Protocol	No Protocol
Nassau (NY)	Envelope	Yes (Multiple)	NA	No Protocol
Philadelphia (PA)	Computer and Geographic Area	Yes (Multiple)	No Protocol	Assigned to Previous Judge
Tarrant (TX)	Computer	Yes (Domestic Violence)	NA	No Protocol
Wayne (MI)	Computer	Yes (Multiple)	Assigned to Previous Judge	No Protocol

*Assignment Mechanisms: The means by which the assignments are made. Computer means assignments are randomly generated by a computer program. Envelope means that judge names are pulled from an envelope for each case. Geographic Area means assignments are based on the region of the case. List means assignments are based on a rotating list.
**Case-type Specialization: Whether the assignment depend in part on the case type? (Case type)
***Repeat Offenders: How are repeat offenders (and probation violators) treated in the procedure?

Table 1b: Courts with As-if Random Assignment Protocols

As-if Random Assignment: Cases are allocated based on one or more of a case's pre-treatment characteristics that are unrelated to potential outcomes.				
County Name (State)	**Assignment Mechanisms***	**Specialization (Type)****	**Repeat Offenders*****	**Pending Case/ Probation*****
Bexar (TX)	Cycled List	Yes (Multiple)	No Protocol	No Protocol
Broward (FL)	Cycled List	Yes (Drug)	No Protocol	Assigned to Previous Judge
Middlesex (MA)	Cycled List	No	NA	NA
New York (NY)	Multi-level List	Yes (Multiple)	No Protocol	Assigned to Previous Judge
Orange (CA)	Cycled List	Yes (Multiple)	Depends	Depends
San Bernardino (CA)	Alpha and Numeric	No	Assigned to Previous Judge	Assigned to Previous Judge
Suffolk (NJ)	Cycled List	Yes (DUI)	Assigned to Previous Judge	NA

*Assignment Mechanisms: The means by which the assignments are made. List means assignments are based on a rotating list. Alpha means assignments are based on the defendant's last name. Numeric means the assignments are based on the case number.
**Case-type Specialization: Whether the assignment depend in part on the case type? (Case type)
***Repeat Offenders: How are repeat offenders (and probation violators) treated in the procedure? Are they assigned to the previous judge?

Table 1c: Courts with Non-random Assignment Protocols

Cases are allocated based on one or more of a case's pre-treatment characteristics that are related to an individual's potential outcomes.				
County Name (State)	**Assignment Mechanisms***	**Specialization (Type)****	**Repeat Offenders*****	**Pending Case/ Probation*****
Alameda (CA)	Calendar	Yes (Multiple)	NA	NA
Bronx (NY)	Calendar	Yes (Multiple)	No Protocol	Assigned to Previous Judge
King (WA)	Geographic Area & Calendar	Yes (Judicial Experience)	No Protocol	No Protocol
Kings (NY)	Geographic Area	Yes (Multiple)	NA	Assigned to Previous Judge
Miami-Dade (FL)	Calendar	Yes (Multiple)	Assigned to Previous Judge	NA
Riverside (CA)	Calendar	Yes (Multiple)	No Protocol	No Protocol
Sacramento (CA)	Calendar	Yes (Domestic Violence)	No Protocol	NA
San Diego (CA)	Geographic Area & Calendar	Yes (Judicial Experience)	No Protocol	No Protocol
Santa Clara (CA)	Geographic Area & Calendar	Yes (Multiple)	No Protocol	Assigned to Previous Judge

*Assignment Mechanisms: The means by which the assignments are made. Calendar means assignments are based off of a rotating or set calendar. Geographic Area means assignments are based on the region of the case.
**Case-type Specialization: Whether the assignment depend in part on the case type? (Case type)
***Repeat Offenders: How are repeat offenders (and probation violators) treated in the procedure?

among the multiple specialized or area-specific judges, allowing for researchers to derive unbiased ATE's for those respective groups. For example, the Philadelphia courthouse has four geographical divisions, each of which has four or five judges. Once a case is assigned into one of those divisions based on the jurisdiction in which the crime was committed, the case is randomly assigned to one of those divisional judges, generally with equal probability.[98] Additionally, most courts do not alter the assignment process for repeat defendants, although a number of them will re-assignment them if they have a pending case or are on probation.

[98] Philadelphia adjusts the probability at which judges are assigned cases based on whether the judge is keeping up with her docket.

ii. Differing Probabilities of Assignment: As was discussed in Subpart III.B, any procedure or event that changes the probability that a given judge or group of judges will be assigned cases requires the researcher to calculate an estimate of the ATE using a re-balanced model (inverse probability weights being the example provided above). The extent to which courts in this study's sample employ procedures that require this rebalancing is mixed (see Table 2, below). Every court consolidates defendants when they are accused of committing the same crime, but only 15 of the 27 courts assign defendants on probation back to their previous judges, and only 4 courts send repeat offenders back to the original judge.

Table 2: Procedures Impacting Probability of Assignment

County Name	Simple or Complete Randomization	Blocking (Type)*	Codefendants Consolidated**	Changes When Overscheduled ***
Alameda (CA)	---	No	NA	NA
Bexar (TX)	---	No	Yes	No
Bronx (NY)	---	Yes (Daily)	Yes	No
Broward (FL)	---	No	Yes	Yes
Clark (NV)	Simple	No	Yes	No
Cook (IL)	Simple	Yes (Daily)	Yes	Yes
Cuyahoga (OH)	Simple	No	Yes	Yes
Dallas (TX)	Simple	Yes (Varies)	Yes	No
Harris (TX)	Complete	No	Yes	No
Hillsborough (FL)	Complete	No	Yes	No
King (WA)	---	No	Yes	No
Kings (NY)	---	No	Yes	Yes
Maricopa (AZ)	Simple	No	Yes	No
Miami-Dade (FL)	---	No	Yes	Yes
Middlesex (MA)	---	No	NA	No
Nassau (NY)	Complete	No	Yes	Yes
New York (NY)	---	No	NA	Yes
Orange (CA)	---	No	Yes	No
Philadelphia (PA)	Simple	No	Yes	Yes
Riverside (CA)	---	No	Yes	No
Sacramento (CA)	---	Yes (Bi-Daily)	Yes	No
San Bernardino (CA)	---	No	No	Yes
San Diego (CA)	---	No	Yes	No
Santa Clara (CA)	---	No	Yes	No
Suffolk (NY)	---	No	Yes	Yes
Tarrant (TX)	Simple	No	Yes	NA
Wayne (MI)	Complex	No	Yes	Yes

*Blocking: Whether cases are assigned in blocks. (On what are the blocks based?)
**Codefendants Consolidated: Whether multiple defendants charged with the same crime combined into one case
***Changes When Overscheduled: Whether cases ever reassigned due to scheduling conflicts.

Roughly half of the courts regularly make changes to the probability of assignment when a judge or group of judges is overscheduled. Some courts merely lower the probability of assignment to a more manageable level, while others take busy judges out of the assignment system entirely. Whether scheduling results in differing probabilities of assignment likely depends on the individual in charge of scheduling and the business judge, meaning that researchers will need to confirm changes in probability using court records and notes.

iii. Post-Assignment Changes: In Subpart III.C we explored the potential for unidentifiable bias due to the ability for judges, administrators, and even parties to seek a change in the original judicial assignment. Every court in the survey has some procedure that allows for post-assignment changes in judicial oversight. These types of changes are most often due to recusal, peremptory challenges or scheduling adjustments, although some courts do make post-assignment changes due to judicial specialization or experience.

iv. Missing Outcomes: Although accounting for attrition is vital to estimating an unbiased ATE, most of the de-randomizing events leading to attrition (particularly missing or lost case data) do not lend themselves to survey-based questions and, as a result, are not represented in the survey data. As was mentioned earlier, however, previous studies have shown substantial potential for attrition due to plea bargains and out-of-court settlements.[99]

v. Additional Insights: While the majority of information gathered from the courts is included in the information above, occasionally a court official or judge would share a thought or say something interesting that fell outside the purview of the survey questions. Additionally, there were a number of general trends that may be informative for researchers hoping to identify courts with usable assignment procedures.

It was clear, for example, that courts are highly concerned with maintaining a perception of unbiased judicial assignment. Before they knew the intentions behind my interest in

[99] *See supra* notes 75 and 76 and accompanying text.

assignment procedures (statistical utilization), nearly every individual I spoke with would start off the discussion by saying something such as, "I can't tell you exactly how assignments are made, but I can assure you that they are definitely random, and have nothing to do with the individuals or issues involved in the case itself." Court officials were initially wary of the survey questions, later divulging that they were worried that I was media looking to write a story on judicial corruption. One individual even told me "we always tell people that these assignments are random all the time, even though there are things that regularly come up that change things around."[100]

While this concern with judicial bias is comforting from an institutional perspective, it poses some difficulty for researchers who are attempting to verify certain assignment procedures. Individuals' insistence on using the word "random" to describe judicial assignment made it tricky to parse out exactly how the assignments were made. Often, surveyors had to ask extremely specific questions or explain the statistical definition for randomness in order to identify the type of system the court used. This confusion is, of course, understandable given the colloquial use of random, but it should serve as a warning for researchers as they communicate with courts.

V. Conclusion

Utilizing procedures that randomly assign judges to cases allows researchers to legitimately address causal questions that would otherwise be intractable. Realizing this, studies taking advantage of this methodology are becoming more and more common. However, these studies often rely on erroneously assumed random assignment and fail to account for all of the de-randomizing events that occur over the life of a case, namely: exceptions to the general assignment procedure, changes in the probability of assignment over the course of the study,

[100] For obvious reasons, this individual asked that I not attribute this quote to him/her or the specific court.

post-assignment changes in judicial oversight, and missing data. Data on 30 state-level criminal courts show that these events are common and should be taken seriously by researchers hoping to harness the advantages of randomly assigned cases.

The conclusions in this Chapter, however, are not meant to dissuade researchers from answering their causal questions using random judicial assignment. Nor do they imply that every study that uses random judicial assignment has to perfectly meet each empirical assumption: researcher-driven experiments, particularly field studies, often feature problems such as imperfect or complicated randomization, non-compliance, spillover, and attrition. But if a study is relying on random assignment it should, at the very least, explore the potential for these issues and, ideally, attempt to account for them using the methodological techniques developed in the experimental literature. Naturally, a more demanding expectation for these studies will mean that some questions cannot be answered using data from certain courts. However, greater methodological rigor will produce more reliable, and therefore, more valuable empirical data.

Finally, this Chapter highlights the importance of both a sound understanding of the methodological specifics of a given empirical technique and a clear comprehension of the rules, procedures, and practices of the court system and the legal institutions more broadly. As a subfield, legal empirical studies are perfectly suited to lead this charge.

Chapter 3[*]

Testing Williams-Yulee*: A Survey Experiment on Judicial Elections, Institutional Trust, and Tenuous Empirical Claims in the Supreme Court*

[*] The experimental study in this Chapter has received IRB approval and was funded by a grant from the Columbia University Graduate School of Arts and Sciences, which is not responsible for the content of this Chapter.

A pre-analysis plan for the experimental design and analysis featured in this Chapter is registered with Evidence in Governance and Politics and can be found at http://egap.org/registration/4706. Any deviations from that pre-analysis plan are noted in the accompanying text.

I. Introduction

In the U.S. Supreme Court's most recent case addressing judicial elections, *Williams-Yulee v. The Florida Bar*,[1] the Court evaluated the constitutionality of a Florida state ban on direct campaign finance solicitations by elected judges. At issue in the case was whether the ban, which stipulated that judicial candidates cannot personally solicit campaign funds but may indirectly secure such funding through their campaign committee, was a violation of the First Amendment.

In ruling on the case, the Supreme Court determined that because the Florida ban was a restriction of the First Amendment's right to free, content-based speech, proper constitutional analysis of the issue required strict scrutiny, the court's most demanding standard of judicial review. In applying this test, the Court found that the ban addressed Florida's "compelling state interest" in "preserving public confidence in the integrity of its judiciary."[2] More contentious, however, was the five-justice majority's conclusion that the restriction was sufficiently "narrowly tailored" to the state's interest. The majority accepted Florida's assertion that a blanket prohibition on direct solicitation was sufficiently tailored because "solicitation by the candidate personally creates a categorically different and more severe risk of undermining public confidence than does solicitation by a campaign committee,"[3] in addition to other testable but empirically unsupported claims. The various dissenting opinions, particularly that of Justice Scalia, highlighted the lack of evidence for this conclusion, while also making a number of similarly unsupported empirical arguments.

With the Supreme Court's stamp of approval, the conclusions regarding public perceptions of judicial legitimacy have subsequently been applied by lower courts in several similar cases. Recently, for example, the 6th Circuit Court upheld Ohio's bans on personal

[1] Williams-Yulee v. The Florida Bar, 135 S. Ct. 1656 (2015).

[2] Williams-Yulee, 135 S. Ct. at 1660.

[3] Williams-Yulee, 135 S. Ct. at 1670.

solicitation by judicial candidates[4] and fundraising 120 days before the primary election or 120 days after the general election. In doing so, the Circuit Court posited that these rules met the narrowly tailored component of strict scrutiny, even though the empirical assumptions required to make that legal conclusion were made "without considering documentary evidence."[5]

In addition to the constitutional questions at play in *Williams*-Yulee, the decision and broader subject matter of the case highlights important and interesting questions about judicial elections and legitimacy that have long been studied by political scientists and legal scholars. How do elected judges balance the opposing responsibilities of judges-as-judges and judges-as-politicians, especially when embedded in a context in which they must—or at least are compelled to—raise campaign funds? And to what extent does this seemingly paradoxical system negatively impact the public's perception of the individual judges, their decisions, and the judiciary as an institution?

This Chapter features a nationally-representative online survey experiment that addresses these broader questions while evaluating the specific empirical arguments made and relied upon in the *Williams-Yulee* opinions, of which only one has been tested using experimental methods.[6] The survey presents subjects with a randomly-assigned hypothetical vignette in which an individual runs for a state trial-level judicial seat and utilizes various

[4] There are three exceptions to Ohio's solicitation rules: 1) A candidate may directly ask for donations when speaking to a crowd of 20 or more individuals, 2) A judge may send personally signed letters asking for donations as long as the donations are directed towards his or her campaign committee, and 3) A judge may send emails asking for donations as long as the donations are directed towards his or her campaign committee (Ohio Code of Judicial Conduct, Cannon 4, Rule 4.4(A).

[5] Platt, et al. v. Bd. of Commissioners of the Ohio Supreme Court, 894 F.3d 235, 243 (6th Cir. 2018) United States Court of Appeals for the Sixth Circuit No. 17-3461 (2018) (ruling that "[t]he district court committed no clear error of judgment in granting the Board's protective order. It determined that fact discovery was unnecessary because such discovery would not aid the court in determining whether Ohio had a compelling interest in maintaining judicial integrity.").

[6] At one point, the majority (the dissent agreed) stated that "judges who personally ask for money may diminish their integrity…" compared to those who do not ask for donations at all (Williams-Yulee, 135 S. Ct. at 1660). Gibson's previous work has demonstrated this using survey experiments (*see* James Gibson et al., *The Effects of Judicial Campaign Activity on the Legitimacy of Courts: A Survey-based Experiment*, 64 POL. RES. Q. 545 (2011)).

campaign fundraising tactics. The survey then presents the subjects with questions relating to the trust and legitimacy they associate with the judge and judicial system presented in the vignette. Contrary to the Supreme Court's assumptions in *Williams-Yulee*, the results of the survey suggest that the U.S. public does not discern a significant difference between direct and indirect judicial solicitation. However, the results indicate that other judicial campaign features (whether the judge has a policy of recusal in cases featuring donors and the amount fundraised) are salient in regard to trust and legitimacy.

This Chapter proceeds in five parts. Part II provides a more detailed factual and legal overview of the *Williams-Yulee* case, investigates the majority and dissenting opinions, and identifies the specific empirical claims that were made and relied upon by the Justices in those opinions. Part III describes the design and implementation of the survey experiment. Part IV analyzes the experimental results as they relate to the *Williams-Yulee* empirical claims, applies the results to several more general hypotheses about judicial elections and public trust, and highlights the inherent limitations of online survey experiments in this context. Part V concludes by briefly discussing what the results mean for *Williams-Yulee* specifically, and might mean for judicial campaign activities more generally, and why it is important for political scientists and legal scholars to test the often-tenuous empirical claims relied upon by the Supreme Court.

II. The Case

A. Factual and Legal Background

The factual background of *Williams-Yulee* is fairly simple.[7] In 2009, Lanell Williams-Yulee, a long-time Florida attorney, decided to run for one of the Hillsborough County Court judicial candidacies that would be filled by popular vote in the 2010 election. As with most

[7] *See* Williams-Yulee, 135 S. Ct. at 1662-64 for case facts.

individuals who run for elected judicial office, Williams-Yulee required funding, and in order to procure donations, she wrote a personal letter announcing her candidacy, highlighting her service and experience, outlining her vision for the office, and asking for financial contributions.[8] The letter was indiscriminately mass-mailed to local voters but, importantly, included her personal salutations and signature and was featured on her campaign website.

Despite her efforts, Williams-Yulee lost the primary election to the incumbent judge. Although she did not take office, she was soon charged with violating Rule 4-8.2(b) of the Rules Regulating the Florida Bar for the manner in which she solicited campaign funds.[9] The prohibition at issue is included in Florida's Code of Judicial Conduct, which was adopted in the 1970's as a response to a number of public judicial scandals involving the Florida Supreme Court.[10] Although the rule prohibited a number of activities related to judicial campaigning and elections, Williams-Yulee was cited specifically for a violation of the ban on direct financial solicitation, which reads:

> A candidate, including an incumbent judge, for a judicial office that is filled by public election between competing candidates shall not personally solicit campaign funds, or solicit attorneys for publicly stated support, but may establish committees of responsible persons to secure and manage the expenditure of funds for the candidate's campaign and to obtain public statements of support for his or her candidacy. Such committees are not prohibited from soliciting campaign contributions and public support from any person or corporation authorized by law.[11]

Williams-Yulee disputed the constitutionality of the law and correctly noted that it prohibited a candidate from sending out letters with her personal signature at the bottom while ostensibly permitting the same exact letter "signed" by her campaign committee, even though

---

[8] For a full text of the letter, see Figure 1: Williams-Yulee Original Letter Text in the appendix for this Chapter.

[9] Fl. St. Bar. Rule 4-8.2(b).

[10] *See* Scott G. Hawkins, *Perspective on Judicial Merit Retention in Florida*, 64 FLA. L. REV. 1421, 1423–1428 (2012).

[11] Code of Judicial Conduct for the State of Florida 6 (2014).

the committee worked for and directly with the candidate. Similarly, under this law, a candidate could not personally approach a friend at a fundraiser and ask for donations, but the candidate's treasurer could make such a direct request, potentially with the candidate standing across the room. Additionally, the Florida Judicial Ethics Advisory Committee had officially interpreted the Code to allow a judicial candidate to serve as the treasurer of her own campaign committee, although the ban disallowed direct solicitation through such a capacity.[12] These observations notwithstanding, the Florida Supreme Court determined that the law was constitutional, and since a violation had occurred, they charged Williams-Yulee for the cost of the Supreme Court proceeding ($1,860) and ordered her to be publicly reprimanded for her actions.[13] Williams-Yulee appealed the Florida Supreme Court ruling to the U.S. Supreme Court, which granted certiorari and eventually heard the case in 2015.

Florida's ban—and its focus specifically on "direct" solicitation—was (and is) far from unique in the United States. At the time that *Williams-Yulee* was heard by the Supreme Court, 39 states held judicial elections at some level, and "nearly every [such] state ha[d] adopted a rule prohibiting judicial candidates from personally soliciting campaign contributions."[14] Many such rules mirror the broad prohibition against personal financial requests implemented in Florida, although some are slightly more or less stringent. Georgia's Canon 7(B)(2),[15] for example, states that candidates for judicial office "shall not themselves solicit campaign funds, or solicit publicly stated support," while allowing the candidate's campaign committee to do so.[16] Similarly, a

---

[12] Florida State Bar Association, An Aid to Understanding Canon 7, pp. 51–58 (2014).

[13] The Fla. Bar v. Williams-Yulee, 138 So. 3d 379 (Fla. 2014), aff'd, Williams-Yulee v. Fla. Bar, 135 S. Ct. 1656 (2015).

[14] Williams-Yulee v. The Florida Bar, 135 S. Ct. 1656 (2015), Petition for Cert, at 2. According to a recent Brennan Center report, 30 of the states that have judicial elections have some sort of ban on solicitation. *See* Alicia Bannon, *Soliciting Donations Discredits the Judiciary,* The Brennan Center (January 26, 2015) at https://www.brennancenter.org/analysis/soliciting-donations-discredits-judiciary).

[15] Georgia Code of Judicial Conduct, Canon 7(B).

[16] As discussed below, this law was deemed unconstitutional by the 11th Circuit in Weaver v. Bonner, 309 F.3d 1312, 1322 (11th Cir. 2002).

judge in Kentucky will be deemed to be engaged in "inappropriate political activity" if he or she directly solicited funding (again, indirect solicitation through campaign committees is appropriate).[17] Punishments for violating these laws vary, with some statutes mandating simple fines, as with the *Williams-Yulee* case, and some allowing for political removal and disbarment.[18]

The question as to whether these bans are constitutionally valid, let alone efficacious, is complex, and no academic or jurisprudential consensus had emerged previous to *Williams-Yulee*.[19] Many saw the laws as an appropriate bulwark against the rising cost and politicization of judicial elections in the United States, arguing that judges, even those who are elected, are tasked with interpreting and following the law and not with directly representing the preferences of their constituents.[20] Consequentially, any interaction between candidate judges and citizen constituents (particularly attorneys that might argue cases in that judges' court) is a serious threat to the "the integrity and independence of the judiciary,"[21] which has been recognized by the Supreme Court as a "vital state interest."[22] The proponents of these bans often have similar reservations regarding any sort of campaign financing by judges—direct or indirect—but cite the increased potential for under-the-table *quid pro quo* transactions between donors and judges and the inherent pressure to avoid retaliation against non-donors as uniquely damaging consequences of a judge personally requesting campaign moneys from individuals.

---

[17] Rules of Supreme Court of Kentucky 3.130(8.2); 4.300, Canon 5. The full clause states, "A judge or a candidate for judicial office shall not solicit campaign funds, but may establish committees of responsible persons to secure and manage the expenditure of funds for the campaign and to obtain public statements of support for the candidacy."

[18] *See, e.g.* Rules of Supreme Court of Kentucky 4.020.

[19] For a recent and particularly thoughtful breakdown of the current debates surrounding elected judges, *see* CHARLES GARDNER GEYH, WHO IS TO JUDGE? 85-98 (2019).

[20] See David E. Pozen, *The Irony of Judicial Elections*, 108 COLUMBIA L. REV. (2008).

[21] Code of Judicial Conduct for the State of Florida 6 (2014), Canon 1.

[22] Caperton v. A. T. Massey Coal Co., 556 U. S. 868, 889 (2009) (internal quotation marks omitted).

In contrast, the Supreme Court has long held that the First Amendment provision preventing Congress from creating laws that abridge the freedom of speech extends to the financing of government elections, and while governments can enact constitutionally sound bans if they are narrowly tailored to serve a compelling interest, such bans are "rare case(s)."[23] Opponents cede that the protection of the court's legitimacy clearly stands as valid interest under such a calculation, but they argue that any law that equates the distribution of mass mailers signed by the candidate with verbal, face-to-face solicitation is either serving some other, non-compelling interest, or has not been sufficiently tailored to be valid under the Constitution.

Although the issue in this case has now been officially resolved by the Supreme Court, the lower courts were starkly divided on whether direct solicitation bans such as Florida's Canon 7C(1) were an unjustified violation of the First Amendment. Of the six circuit courts that had reviewed such laws previous to *Williams-Yulee*, four determined that they did not pass constitutional muster. The Eleventh Circuit invalidated the Georgia ban discussed above on two fronts, finding that such a broad ban is over inclusive by "completely chill[ing] a candidate's speech" while "hardly advancing the state's interest in judicial impartiality at all."[24] Similarly, the Eighth Circuit did not believe that the distinction between a letter signed by a candidate and one sent from that candidate's campaign constituted narrow tailoring,[25] and the Ninth Circuit suggested that such behavior "presents little to no risk of corruption or bias towards future litigants…."[26]

The Third and Seventh Circuits, on the other hand (along with the Supreme Courts of

[23] *Burson* v. *Freeman*, 504 U. S. 191, 211 (1992) (plurality opinion).

[24] Weaver v. Bonner, 309 F.3d 1312, 1323 (11th Cir. 2002).

[25] *See* Republican Party of Minnesota v. White, 416 F.3d 738 (8th Cir. 2005) (en banc) (White II).

[26] Wolfson v. Concannon, 750 F. 3d 1145, 1157 (9th Cir. 2014).

Arkansas,[27] Florida,[28] and Oregon[29]) found that nearly identical bans did indeed meet the compelling interest requirement while not being so overly broad that they were not narrowly tailored. In *Stretton v. Disciplinary Board of Supreme Court of Pennsylvania*,[30] a judicial candidate in Pennsylvania was barred from sending campaign letters bearing his signature, and the Seventh Circuit upheld direct solicitation bans in both Wisconsin and Indiana that had been challenged by a pair of sitting judges.[31] These conclusions—that the bans were in service of a compelling interest and narrowly tailored to effectuate such an interest—were ultimately portents for the determination that the Supreme Court would make in *Williams-Yulee*.

B. The Supreme Court Decision

In an opinion penned by Chief Justice Roberts and joined by Justices Breyer, Sotomayor, Kagan, and Ginsburg (Ginsburg did not join in Section II of this opinion, writing her own opinion), the majority upheld the Florida ban on the grounds that it was narrowly tailored to Florida's interest in "protecting the integrity of the judiciary" and "maintaining the public's confidence in an impartial judiciary."[32] Specifically, Roberts relied on a determination made in *Caperton v. A.T. Massey Coal Co.*[33] (arguably the Court's most prominent case dealing with judicial campaign finance) that the public's perception of the judicial institution is of paramount concern to state legislators and citizens.

The dissent agreed with the majority that states do have a valid interest in maintaining legitimacy, and they endorsed the strict—more specifically, exacting—scrutiny that the majority used in determining whether the law was an appropriate means for achieving such an interest,

---

[27] *See* Simes v. Arkansas Judicial Discipline and Disability Commission, 247 S.W.3d 876 (Ark. 2007).

[28] Williams-Yulee, 135 S. Ct. at 1663.

[29] In re Fadeley, 802 P.2d 31 (Or. 1990) (per curiam).

[30] Stretton v. Disciplinary Board of Supreme Court of Pennsylvania, 944 F.2d 137 (3d Cir. 1991).

[31] *See* Bauer v. Shepard, 620 F.3d 704 (7th Cir.2010) (the Wisconsin case) and Siefert v. Alexander, 608 F.3d 974 (7th Cir.2010) (the Indiana case).

[32] The Fla. Bar v. Williams-Yulee, 138 So. 3d at 385.

[33] Caperton v. A.T. Massey Coal Co., 556 U.S. 868.

but they did not agree with the majority's conclusion that the Florida ban on direct solicitation was narrowly tailored to serve that end. Williams-Yulee had argued this point based on two divergent assertions. First, that the ban was too narrow. Because the law failed to prohibit other, equally problematic campaign behavior such as direct solicitation by campaign committees and the writing of personal (and direct, under the definitions adopted by the majority) thank you letters, it appeared insincere on its face and may be detrimental to the state's interest in legitimacy. Second, Williams-Yulee claimed that the law was too broad, banning a set of activities that are essential to the election process and not threats to judicial legitimacy.

In addressing the underinclusiveness claim, the majority cited a number of cases where the Court had upheld laws under strict scrutiny even when they did "not address all aspects of a problem in one fell swoop."[34] In making a distinction between the case at hand and the cases in which underinclusiveness did preclude narrow tailoring,[35] Roberts argued that the Florida ban was focused "squarely at the conduct most likely to undermine public confidence in the integrity of the judiciary: personal requests for money by judges and judicial candidates,"[36] and that direct solicitation by campaign committees is categorically different, thereby creating less risk of delegitimizing the judicial institution. Similarly, the majority concludes that thank you letters "may be" detrimental to the state's interests, but that the risk that such behavior imposes is less than the risk of direct solicitation.

C. Empirical Claims

In supporting their legal conclusions, both the majority and the dissent presented a number of claims that were not empirically supported in the opinions or the materials provided

[34] Williams-Yulee, 135 S. Ct. at 1668 (citing as examples, *Burson*, 504 U. S., at 207; see *McConnell*, 540 U. S., at 207–208; *Metromedia, Inc.* v. *San Diego*, 453 U. S. 490, 511–512 (1981) (plurality opinion); *Buckley*, 424 U. S., at 105).

[35] *See* Church of Lukumi Babalu Aye, Inc. v. Hialeah, 508 U. S. 520, 543– 547 (1993); Smith v. Daily Mail Publishing Co., 443 U. S. 97, 104–105 (1979).

[36] Williams-Yulee, 135 S. Ct. at 1668.

to the courts by the parties.[37] While making claims without providing substantial proof is certainly not a novel practice for the court[38]—indeed, it is often a necessity given the fact that they must make determinations using the evidence provided to them—the assumptions relied on in *Williams-Yulee* played a particularly pivotal role in the legal conclusions that the justices drew and ultimately shaped the outcome and jurisprudential effect of the case. As a result, the legal foundation of the case depends on the empirical validity of the claims.

Empirical Claim 1: Candidates for judicial office who are actively raising campaign funds will diminish their own integrity and produce a lower sense of institutional trust and legitimacy than candidates who are not actively raising campaign funds.

While this claim is not explicitly key to the decision in *Williams-Yulee* (the majority and dissent agree that solicitating campaign funds is detrimental to public trust and legitimacy relative to non-solicitation) and has been verified by previous research,[39] the claim is important to explore because it presupposes the conclusions drawn by the majority. If the active pursuit of campaign financing by judicial candidates does nothing to impact the public's perception of those individuals, or the judicial office more broadly, any prohibitions on solicitation (whether direct or indirect) should not pass strict scrutiny, as the banned activities would be unrelated to the state's stated interest in maintaining judicial legitimacy. Additionally, in providing support against the dissent's claims of underexclusiveness, the majority suggested that direct solicitation is the campaign behavior most likely to damage the institutional reputation of the courts.

---

[37] While some of the amici filed in the case cited survey research touching on the public perception of judicial solicitation, none of the data directly addressed the relative comparisons the Justices made and relied upon in the case. *See, e.g.,* Amicus Brief of The Brennan Center in Support of Appellees, Williams-Yulee v. The Florida Bar, 135 S. Ct. 1656 (2015).

[38] *See e.g.,* Adam Liptak, *Did the Supreme Court Base a Ruling on a Myth?*, N.Y. Times, March 6, 2017.

[39] *See* James L. Gibson, *Challenges to the Impartiality of State Supreme Courts: Legitimacy Theory and "New-Style" Judicial Campaigns*, 102 AM. POL. SCI. REV. 59-75 (2008); James L. Gibson & Gregory A. Caldiera, *Campaign Support, Conflicts of Interest, and Judicial Impartiality: Can Recusals Rescue the Legitimacy of Courts?*, 74 *J. OF POL.* 18-34 (2012); James L. Gibson & Gregory A. Caldiera, *Judicial Impartiality, Campaign Contributions, and Recusals: Results from a National Survey*, 10 *J. EMP. LEGAL STUD.* 76-103 (2013); James L. Gibson, et al., *The Effects of Judicial Campaign Activity on the Legitimacy of Courts: A Survey-based Experiment*, 64 POL. RES. Q. 545-558 (2011).

Empirical Claim 2: Candidates for judicial office who directly solicit campaign funds will diminish their own integrity and produce a lower sense of institutional trust and legitimacy than candidates who raise campaign funds indirectly through campaign committees.

This is the crucial empirical claim in the case, with the majority accepting Florida's contention that "solicitation by the candidate personally creates a categorically different and more severe risk of undermining public confidence than does solicitation by a campaign committee,"[40] and the dissent arguing that this "intuition"[41] does not meet the necessary burden required to pass the second half of the strict scrutiny test.

Empirical Claim 3: Candidates for judicial office who send thank you notes to campaign donors will not diminish their own integrity more or produce a lower sense of institutional trust and legitimacy than candidates who do not send thank you letters.

Williams-Yulee also argued that a sufficiently tailored state action would also ban personal thank you letters to donors (which the Florida law does not do), because doing so would ostensibly weaken the public's trust in the judiciary by furthering the financial connection between the candidate and donor through direct communication. The majority disagreed, positing that "[p]ermitting a judicial candidate to write thank you notes to campaign donors likewise does not detract from the State's interest in preserving public confidence in the integrity of the judiciary.[42] Evidence supporting this argument would invalidate Florida's ban under the underinclusivity prong of the narrowly tailored test, which recognizes that underinclusive laws may raise "doubts about whether the government is in fact pursuing the interest it invokes, rather than disfavoring a particular speaker or viewpoint."[43]

Empirical Claim 4: Attorneys (and citizens generally) are more pressured by and therefore more likely to "donate" to a judicial candidate when that candidate directly solicits campaign funds than if the candidate solicited campaign funds indirectly through her campaign committee.

---

[40] Williams-Yulee, 135 S. Ct. at 1669.

[41] Williams-Yulee v. The Florida Bar, 135 S. Ct. 1656 (2015), Dissent (Scalia) at 1678 (quotations omitted).

[42] Williams-Yulee, 135 S. Ct. at 1660.

[43] Williams-Yulee, 135 S. Ct. at 1668 (quoting *Brown v. Entertainment Merchants Ass'n*, 131 S. Ct. 2729 at 2740 (2011).

The majority also argued that direct solicitation "inevitably creates pressure for the recipient to comply, in a way that solicitation by a third party does not,"[44] focusing specifically on the fact that a significant portion of campaign donations come from and are solicited toward attorneys. While this claim does not serve as the lynchpin for the majority's conclusion, disproving it empirically would weaken its overall argument.

III. The Survey Experiment

A. Methodology and Venue

This Chapter utilizes an online survey experiment to measure public response to the various campaign fundraising approaches scrutinized by the Supreme Court Justices in *Williams-Yulee*. As with standard public opinion surveys, survey experiments present subjects with information and ask a number of related questions in hopes of gauging the subjects' opinions, cognitive ability, or behavior. Unlike regular surveys, however, survey experiments randomly vary either the information provided to the subjects or vary the questions asked in order to identify the causal effect of such variations.

The methodological value of randomized experiments is well known.[45] When designed and conducted properly, randomized experiments—and survey experiments by extension—allow researchers to overcome the fundamental problem of causal inference by creating two or more comparison groups that are, in expectation, statistically equivalent except for the existence or nature of the randomly-assigned treatment.[46] The researcher can then compare the average outcomes of the groups, knowing that any difference between them is due to the causal effect of

---

[44] Williams-Yulee, 135 S. Ct. at 1669.

[45] *See* Donald B. Rubin, *Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies*, 6(5) *J. EDU. PSY.* 688-701 (1974).

[46] This problem states that you can never truly know the effect of a treatment on a specific object because you can never see two outcomes for one given case. *See* Paul W. Holland, *Statistics and Causal Inference*, 81(396) *J. AM. STAT. ASS'N* 945-60 (1986).

the treatment.[47] In the survey featured in this Chapter, the experimental treatments are embedded in a hypothetical vignette based closely on the letter used by Williams-Yulee in her campaign.

The surveys were distributed using Lucid's Fulcrum Exchange. Lucid is a global online sampling company that provides demographic-specific survey populations for market research and academic studies. A recent, independent study showed that Lucid's survey respondents are nearly as demographically representative as traditional national probability samples and produce equivalent political, psychological, and experimental results.[48] This same study also compared the experimental results of the Lucid sample against one collected through Amazon's Mechanical Turk and found that it provides a more externally valid (albeit more expensive) set of respondents. For the survey in this Chapter, Lucid provided 1,038 respondents who were quota-sampled to reflect national U.S. demographics, and pulled from Florida (33% of the sample) and the other 38 states that feature judicial elections (67% of the sample). The descriptive statistics of this sample are discussed in *infra* Subpart IV.A.

B. Survey Design and Treatments

The survey is comprised of three main sections. The first section includes demographic questions asking the subjects' age, sex, education, income, race, state of residence, political ideology, political party affiliation, and previous presidential vote. This section also includes two background questions asking whether the subjects have been personally involved in a court case and whether the subjects personally know any judge or former judge.[49] The second section presents the subjects with a three-part hypothetical vignette in which the reader is given a short

[47] *See* Donald B. Rubin, *Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies*, 6(5) *J. EDU. PSY.* 688-701 (1974).

[48] Alexander Coppock and Oliver A. McClellan, *Validating the Demographic, Political, Psychological, and Experimental Results Obtained from a New Source of Online Survey Respondents*, 6 RESEARCH & POL. 1-14 (2019).

[49] *See infra* Appendix: Survey Text and Logic, Blocks 2-3.

description of an individual named George Anderson who is running for an open trial-level judicial seat in Fairview County (a fictional county), Florida.[50] The third section then asks the subjects questions about the vignette to gauge how much the circumstances surrounding Anderson's campaign affected their perceptions of Anderson as a fair and impartial judge, and the legitimacy they afford to his future judicial decisions and the court in which he sits(the specific outcome measures used in the survey are outlined and justified in *infra* Subpart III.D).[51]

The hypothetical vignette contains the experimental treatments, and by extension, is the mechanism through which the effect of various judicial fundraising and election tactics can be estimated. Part 1 of the vignette introduces George Anderson and details the actions he took leading up to the election. Part 1 varies on the following experimental factors:

i. Active Fundraising v. No Fundraising:[52] One of the foundational empirical assumptions made by both the majority and dissenting Justices in *Williams-Yulee* is that an electoral system in which judicial candidates actively solicit campaign funds result in less trust and institutional legitimacy. To test this claim, one of the variations of the vignette (the "control" condition) describes Anderson's entry into the judicial race, and recognizes the necessity of raising campaign funds, but explains that Anderson did not directly solicit donations. This variation begins as follows:

> *Last year, Fairview County (a fictional county in Florida) held elections for the judges in their trial courts. A citizen and attorney of Fairview County named George Anderson decided that he would like to run for one of the judicial positions. Like most of the judicial candidates running in the election, George Anderson needed to raise money for his campaign.*
>
> *In order to do this, Anderson included his name on the official election registry. While he did not solicit or ask for any donations, he was willing to take donations*

[50] *See infra* Appendix: Survey Text and Logic, Blocks 4-7.

[51] *See infra* Appendix: Survey Text and Logic, Blocks 8-14.

[52] The active fundraising treatment was embedded into versions B-J of the vignette, and the no-fundraising treatment was embedded into version A of the vignette (*see infra* Appendix B: Survey Text and Logic, Blocks 4-7).

from individuals who knew he was running for office and wanted to support his campaign.

This text is followed by a statement by Anderson that the subjects were told appeared on his website. The statement is a slightly altered version of the campaign letter used by Williams-Yulee in her bid for judicial office,[53] and reads as follows:

Dear Friends:

I have served as a public servant for this community as public defender as well as a prosecutor for the past 18 years. Having been involved in various civic activities such as "The Great American Teach In", Inns of Court, pro bono legal assistance, Metropolitan Ministries outreach program, as well as a mentor for various young men and women residing within Fairview, I have long worked for positive change in this county and in Florida broadly.

With the support of my family, I now feel that the time has come for me to seek elected office. I want to bring fresh ideas and positive solutions to the Fairview County judicial branch. I am certain that I can uphold the laws, statutes, and ordinances as prescribed by the Constitution of The State of Florida as well as the Constitution of the United States Of America.

Thank you in advance for your support!

George Anderson

The other nine variations on the vignette have a similar introduction and message, but subjects are told that Anderson (or his campaign) either sent letters explicitly asking for donations or organized a campaign donation event. The campaign letter in the solicitation vignettes also concluded with the following appeal:

To succeed in this effort, I need to mount an aggressive campaign. I'm inviting the people that know me best to join my campaign and help make a real difference. An early contribution of $25, $50, $100, $250, or $500, made payable to the "George Anderson Campaign for County Judge" will help raise the initial funds needed to launch the campaign and get out our message to the public.

[53] To compare to Williams-Yulee's original letter, *see infra* Appendix B: Text of Original Williams-Yulee Letter.

ii. Direct Solicitation v. Indirect Solicitation:[54] The key empirical claim in *Williams-Yulee* was that judges directly asking for funding (whether in person or through letters) posed a greater threat to Florida's interest in judicial trust and legitimacy than judges who fundraised indirectly through their campaigns. To test this claim, subjects who received the direct solicitation vignettes were either told that Anderson himself sent and signed the campaign letters asking for donations ("*Anderson sent the following letter . . .*") or that he organized and attended a campaign donation event, where he personally asked for funding ("*Anderson organized a campaign donation event, at which he personally asked guests . . . to donate to his campaign . . .*"). Conversely, those assigned to the indirect vignettes, were told that the Anderson's campaign sent and signed the letters ("*Anderson had his campaign committee send the following letter . . .*") or that it was his campaign committee that asked for donations at the campaign event ("*Anderson organized a campaign donation event (Anderson was not in attendance), at which his campaign committee asked guests . . . to donate to his campaign . . .*").

iii. Citizen Donors v. Attorney Donors:[55] In their opinions, the Justices also highlighted the particularly worrisome relationship between judicial candidates and attorney donors who would potentially argue cases in the judge's courtroom. In some of the vignettes, subjects were told that Anderson's campaign efforts were focused on the citizens of Fairview County, while other vignettes described Anderson or his campaign reaching out to attorneys.

iv. Letter Solicitation v. Face-to-face Solicitation:[56] The dissent in *Williams-Yulee* argued that Florida's ban on direct solicitation was not narrowly tailored because it did not draw

---

[54] The direct solicitation treatment was embedded into versions B, C, F, and G of the vignette, and the indirect solicitation treatment was embedded into versions D, E, H, and I of the vignette (*see infra* Appendix B: Survey Text and Logic, Blocks 4-7).

[55] The citizen treatment was embedded into versions C, E, G, I of the vignette, and the attorney treatment was embedded into versions B, D, F, H of the vignette (*see infra* Appendix B: Survey Text and Logic, Blocks 4-7).

[56] The letter treatment was embedded into versions B-E of the vignette, and the attorney treatment was embedded into versions F-I of the vignette (*see infra* Appendix B: Survey Text and Logic, Blocks 4-7).

a distinction between writing mass mailings and making face-to-face solicitations, the latter of which they argued is more detrimental to judicial trust and legitimacy. To test this, some subjects were told that in order to raise funds, Anderson or his committee sent a campaign letter, while others were told that Anderson organized a campaign donation event, where either he or his campaign committee asked for donations.

Parts 2 and 3 of the vignette provide some additional information regarding the outcome of Anderson's fundraising efforts and how he planned to approach the potential conflicts of interest that stem from the procured donations.[57] Some of the treatment variations included in these final parts do not directly address arguments raised in *Williams-Yulee* but, as is briefly discussed in Subpart IV.D, these treatments are valuable in addressing other questions relating to judicial elections more broadly

v. Varying Success in Fundraising:[58] The amount of trust and legitimacy that subjects attribute to a given campaign strategy could depend on the amount of money that results from that strategy. In a short paragraph in part 2 of the vignette, subjects were given either an average or total amount raised and were told that the average/total amount was either high, medium, or low. "*The **average/total** donation amount resulting from these efforts was **$1,010 per donor/$480 per donor/ $45 per donor//$4,500 / $48,000/$101,000.***" [Bold sections were randomized].

vi. Thank You Letter v. No Thank You Letter:[59] In her brief, Williams-Yulee argued that the ban was not narrowly tailored because it allowed judges to send personal thank you letters to

[57] These final two sections were functionally distinct from the main body of the vignette in the survey software, so they could be randomly assigned independently of the 10 treatment variations in part 1 (*see infra* Appendix B: Survey Text and Logic, Blocks 6-7).

[58] The average treatment was embedded into versions K-M of part 2 of the vignette, the total treatment was embedded into versions N-P. The low amount treatment was embedded into versions M and P of the vignette, the medium amount treatment was embedded into versions L and O of the vignette, and the high amount treatment was embedded into versions K and N of the vignette (*see infra* Appendix B: Survey Text and Logic, Block 6).

[59] The thank you letter treatment was embedded into versions T-V of part 3 of the vignette, the no thank you letter treatment was embedded into versions Q-S of part 3 of the vignette (*see infra* Appendix B: Survey

donors, which she argued is just as determinantal to trust and legitimacy as direct solicitation. The majority did not agree. To test this, some subjects were told in part 2 of the vignette that "*To express his gratitude for those who helped him win his election, Anderson sent personally-signed thank you letters to every individual who donated to his campaign.*"

vii. Recusal v. Non-recusal:[60] The negative impact that campaign financing has on perceptions of trust and legitimacy could be tied to whether donations will lead to potential conflicts of interest in the courtroom.[61] In the final section of the vignette, subjects were told that Anderson ended up winning the election and that "*[w]hen he later realized that some of those that donated to his campaign would end up as parties or attorneys in his court*," he either "*decided to remove himself from any cases involving a donor (the cases were reassigned to other judges)*" or "*decided not to remove himself from any cases involving a donor (the cases were not reassigned to other judges).*" To test whether the explicit discussion of recusal alone impacts levels of trust and legitimacy, some subjects were simply told that Anderson ended up winning the election, without any reference to recusal.

C. Randomization

Subjects were randomly assigned within a given treatment category independently of their assignment to the other categories. As described above, each subject was presented with a three-part vignette: the first section included one of ten variations of the campaign strategy used by the judge; the second section included one of six variations of the donation results; and the third section included one of six variations of the post-election behavior (recusal and thank you letters). The probability of assignment to the variations of the campaign strategy were a 2/11

Text and Logic, Block 7).

[60] The recusal treatment was embedded into versions Q and T of part 3 of the vignette, the no recusal treatment was embedded into versions R and U of part 3 of the vignette, and the non-discussion treatment was embedded into versions S and V of the vignette (*see infra* Appendix B: Survey Text and Logic, Block 7).

[61] *See* Chapter 1 of this dissertation for a detailed discussion of judicial recusal as it relates to campaign finance.

chance of being assigned to the control (no fundraising) condition and a 1/11 chance of being assigned to the remaining nine conditions. The probability of assignment for the second and third sections of the vignette were equal for all variations (1/6).

When all three parts are combined, the result is 10 x 6 x 6 = 360 unique vignettes. While this creates a number of unique treatments that is over 33 percent of the total subjects in the study, which would normally be statistically underpowered (to say the least), by using a factorial randomization scheme where the probability of being assigned one aspect of the treatment is independent of being assigned another aspect of the treatment, the survey responses to distinct vignettes can be averaged to create groups that allow for sufficiently powered testing.

D. Outcomes

The survey features three primary outcomes that measure levels of judicial trust, judicial fairness, and institutional legitimacy.[62] While the outcomes are not perfect measures of each concept, they are each adopted from the measures used by the leading political scientists and legal scholars in the field of judicial legitimacy.[63] After the survey presented them with the randomized vignette, each survey subject was presented with the following three questions. The order of the first two were randomized, and the third question contains a randomized element that was included as part of a separate study recusal. Although answer selections to the judicial bias question indicate increasingly positive perceptions,[64] the numerical values of these answers were recoded to reflect the increasingly negative answer options in the fairness and legitimacy questions.

---

[62] In addition to these three outcome measures, the survey also included an additional outcome measure on legitimacy that is used as a measure for survey consistency. This outcome is not reported as part of the experimental analysis, pursuant to this Chapter's pre-analysis plan.

[63] *See, e.g.,* Tom R. Tyler, *Psychological Perspectives on Legitimacy and Legitimation*, 57 *ANNUAL REVIEW OF PSYCHOLOGY* 375-400 (2006); TOM R. TYLER & YUEN J. HUO, TRUST IN THE LAW: ENCOURAGING PUBLIC COOPERATION WITH THE POLICE AND COURTS (2002). For examples of research focused on determining the perceived fairness and impartiality of judges and their decisions, see the works of Gibson and Caldeira, *supra* note 39.

[64] This was done to vary the answer selections across outcome measures to help ensure more thoughtful responses by the subjects.

i. Judicial Fairness and Impartiality:

Do you believe Judge Anderson can serve as a fair and impartial judge for Fairview County?
- *I strongly believe Anderson can be fair and impartial*
- *I somewhat believe Anderson can be fair and impartial*
- *I somewhat believe Anderson cannot be fair and impartial*
- *I strongly believe Anderson cannot be fair and impartial*

ii. Institutional Legitimacy of the Court System:

Assume for the moment that all elected judges in the Fairview County Court used the same campaigning strategies as Judge Anderson. How legitimate would you consider the Fairview County Court?
- *I would consider it a very legitimate institution*
- *I would consider it a somewhat legitimate institution*
- *I would not consider it a very legitimate institution*
- *I would not consider it a very legitimate institution at all*

iii. Judicial Bias:

Assume for a moment that one of the individuals who donated to Judge Anderson is now ***[an attorney who is arguing / a party in]*** *a case in Judge Anderson's Court. Do you think that the relationship between the donor and Judge Anderson will bias Judge Anderson's judicial behavior? [Bolded section is randomly assigned]*
- *It will definitely bias the judge's behavior*
- *It will probably bias the judge's behavior*
- *It will probably not bias the judge's behavior*
- *It will definitely not bias the judge's behavior*

Additionally, the survey also asked the subjects how willing they would be to donate to Anderson's campaign. The specific wording of this question was dependent on the treatments assigned to the subject. The donation outcome is used to test Empirical Claim 4.

iv. Likelihood of Donating:[65]

Imagine that you were an attorney in Fairview County during the judicial election. You receive the letter from Anderson asking for donations. How likely would you be to donate to his campaign?
- *Extremely likely*
- *Moderately likely*
- *Slightly likely*

[65] This version of the donation question appeared for subjects who were assigned to the treatment conditions in which Anderson sent letters asking for support.

- *Neither likely nor unlikely*
- *Slightly unlikely*
- *Moderately unlikely*
- *Extremely unlikely*

IV. Analysis and Results

A. Survey Sample:

An initial sample of 100 subjects were surveyed on June 18, 2018. After verifying that the survey was functioning correctly, the remaining 938 subjects were surveyed from June 20 to June 22. The resulting final sample consisted of 1,038 subjects recruited from the 33 states in

Table 1a: Sample Demographics

Respondent Characteristic	Mean	Median	Minimum	Maximum	Proportion of Missing Responses
Age (Years)	43.19	41	18	92	79/1,038 (7.6%)
Sex (Female)	0.54	Female	NA	NA	1/1,038 (0%)
Education (College or Higher)	0.43	Some College	8th Grade	Post-graduate Degree	0/1,038 (0%)
Income (US Dollars)	$56,898	$49,999	$0	$350,000	64/1,038 (6.2%)
Race (White)	0.76	White	NA	NA	0/1,038 (0%)
Ideology (5-point Scale; Low = Conservative)	2.94	3	1	5	1/1,038 (0%)
Political Party Affiliation (Democrat)	0.39	Republican	NA	NA	2/1,038 (0%)
Vote in Last Presidential Election (Trump)	0.35	Clinton	NA	NA	0/1,038 (0%)
Personal Experience in Court (Yes)	0.27	No Experience	NA	NA	1/1,038 (0%)
Personally Know a Judge (Yes)	0.14	Do Not Know Judge	NA	NA	1/1,038 (0%)

which elections are used for the initial selection of trial or appellate judges.[66] To independently test the claims in *Williams-Yulee* from the state in which the case originated, Lucid oversampled from Florida (33% of the sample). The remaining subjects were sampled from the remaining judicial election states according to the proportional size of each state's population. Sample demographics generally reflect the demographic distribution of the United States, although the sample is slightly over-educated, white, and wealthy (see Table 1a, below).

B. Survey Robustness Checks:

Using hypothetical vignettes as a mechanism for testing empirical claims is common practice within social science disciplines, and the details of the vignette used in this survey closely match the circumstances at issue in *Williams-Yulee*. But the campaign strategies embedded in the survey treatments will almost certainly be less salient than they would be if subjects interacted with them in the real world (i.e. they were residents of Florida and had heard about Williams-Yulee's real campaign). Thus, a null finding for any of the hypotheses tested may be due to either the incorrectness of the empirical assumptions made by the Justices in *Williams-Yulee* being, or simply due to limitations in the experimental design.

To strengthen the validity of the study, a "extreme" treatment variation was introduced that described clearly inappropriate *quid pro quo* campaign behavior that would almost certainly impact how a judge and the judiciary as a whole was viewed by the public. Roughly ten percent of the subjects were given a vignette in which the campaign behavior of Anderson was the following:

> *In order to do this, Anderson directly asked a number of prominent attorneys in Fairview County for campaign money, promising to "make it up to them" when*

[66] These states are: Alabama, Arizona, Arkansas, California, Florida, Georgia, Idaho, Illinois, Indiana, Kansas, Kentucky, Louisiana, Maryland, Michigan, Minnesota, Mississippi, Missouri, Montana, Nevada, New Mexico, New York, North Carolina, North Dakota, Ohio, Oklahoma, Oregon, Pennsylvania, South Dakota, Tennessee, Texas, Washington, West Virginia, Wisconsin. States that do not hold any judicial elections (Delaware, Hawaii, Massachusetts, New Hampshire, New Jersey, Rhode Island, and Virginia), only hold retention elections (Alaska, Colorado, Iowa, Nebraska, Utah, and Wyoming), or only hold elections for probate court judges or magistrate judges (Connecticut, Maine, South Carolina, and Vermont) were not included in this sampling.

> *they argued cases in his courtroom. While he did not guarantee that they would win those cases, he did promise to give extra considerations to their arguments if they made significant donations to his campaign.*[67]

A null finding for this treatment will thus indicate that either the treatment variations are not salient in the structure of the survey vignette, or that the survey respondents are simply not reading the vignette carefully.

Table 2a shows that the *quid pro quo* treatment caused large and statistically significant differences in all three of the primary outcomes of interest. Subjects who were presented with a vignette that included the judge explicitly promising better case outcomes for donors are more than 1.1 points less likely to see the judge as fair and impartial compared to subjects presented with a hypothetical judge who solicited contributions through more the more proper tactics discussed previously. The differences also apply to perceptions of institutional legitimacy (1.1 points) and to a lesser extent, perceptions of bias (.47 points). On a four-point scale, in which the maximum movement is three points, these differences provide strong evidence that the structure of the experiment and the attentiveness of the subjects allow the subjects to be "treated" by the vignettes.

Interestingly, the *quid pro quo* treatment did not produce a significant effect on the likelihood that a subject would donate to Anderson's campaign, although the direction of the estimate is in the expected direction. This suggests that the donation measurement may not be reliable even in the context of the survey, in addition to being suspect as a measurement for actual behavior.

[67] The *quid pro quo* treatment was embedded into versions J of the vignette (*see infra* Appendix B: Survey Text and Logic, Blocks 2-4).

Table 2a: Regular Solicitation v. Quid Pro Quo

Outcome	Mean of Regular Solicitation (n = 942)	Mean of Quid Pro Quo (n = 96)	Estimated ATE (Regression w/ Covariates**)	P-value* (2-tailed)	Lower 95% Confidence Interval	Upper 95% Confidence Interval
Fair and Impartial Judge	1.975	3.011	**1.115**	**0.000**	1.018	1.213
Legitimate Court	2.079	3.105	**1.105**	**0.000**	1.004	1.206
Bias Due To Donation	2.764	3.189	**0.465**	**0.000**	0.374	0.557
Likelihood of Donating	3.739	3.702	-0.26	0.860	-0.172	0.121

* When accounting for the ten separate tests run on each of the three outcome measurements, the Šidák correction results in an alpha level of .005. **Coefficients and p-values in bold fall below the Šidák-corrected significance level.**
** Pre-treatment covariates include: age, sex, education, income, race, state of residence, political ideology, political party affiliation, previous presidential vote, whether the subjects have been personally involved in a court case, and whether the subjects personally know any judge or former judge.

To further test the extent to which subjects were engaging with the survey language, one additional outcome measure for impartiality not in the experimental analysis was provided in the survey:[68]

As a citizen, how likely would you be to accept decisions made by Judge Anderson as impartial, fair, and legitimate?

- *Extremely likely*
- *Moderately likely*
- *Slightly likely*
- *Neither likely nor unlikely*
- *Slightly unlikely*
- *Moderately unlikely*
- *Extremely unlikely*

This question was presented on a separate page from the impartiality question included in the experimental analysis but should be highly correlated if the subjects are reading the

[68] This additional outcome was designated as a consistency-check variable in this Chapter's pre-analysis plan and was not designated for experimental analysis.

survey and answering questions consistently. The Chronbach's Alpha between these two impartiality measures was 0.88, suggesting strong consistency across questions in the survey.

C. Testing the Empirical Claims in *Williams-Yulee*:

To test each of the empirical claims asserted by the Supreme Court Justices in *Williams-Yulee*, the survey responses were compared to three outcome measures—fairness and impartiality, institutional legitimacy, and likelihood of bias (e.g., willingness to donate for the donation claim). The comparison across treatment groups employs multivariate regression analysis that includes each of the demographic variables collected in the first part of the survey, with the exclusion of state of residence.[69] P-values from the regression were subjected to a Šidák Correction for multiple comparisons.[70]

Empirical Claim 1 posited that candidates for judicial office who are actively raising campaign funds will diminish their own integrity and produce a lower sense of institutional trust and legitimacy than candidates who are not actively raising campaign funds. The claim was tested by comparing the outcome responses of subjects who were presented with the no-solicitation control vignette with those who were presented with the vignettes in which Anderson did activity seek donations, either personally or through his committee. The results of the multivariate regression of solicitation on the outcome measures (see Table 3a) suggest that solicitation does have a slight negative impact on perceptions of impartiality, legitimacy, and bias (higher values in each outcome measure indicate more "negative" perceptions), but the estimated effects are not significant.

[69] The categorical demographic variables were re-coded to binary variables as follows: Sex (1 = Female), Education (1 = College graduate or higher), Race (1 = White), Party affiliation (1 = Democrat), Presidential vote (1 = Trump).

[70] Zbyněk Šidák, *Rectangular Confidence Regions for the Means of Multivariate Normal Distributions*, 62 J. AM. STAT. ASS'N 626-633 (1967).

Table 3a: Non-Solicitation v. Solicitation

Outcome	Mean of Non-Solicitation (n = 187)	Mean of Solicitation (n = 755)	Estimated ATE (Regression w/ Covariates**)	p-value (2-tailed)	Lower 95% Confidence Interval	Upper 95% Confidence Interval
Fair and Impartial Judge	1.886	1.997	0.118	0.447	-0.018	0.135
Legitimate Court	2.032	2.090	0.058	0.106	0.045	0.191
Bias Due To Donation	2.716	2.776	0.086	0.221	0.016	0.156

* When accounting for the ten separate tests run on each of the three outcome measurements, the Šidák correction results in an alpha level of .005. **Coefficients and p-values in bold fall below the Šidák-corrected significance level.**

** Pre-treatment covariates include: age, sex, education, income, race, state of residence, political ideology, political party affiliation, previous presidential vote, whether the subjects have been personally involved in a court case, and whether the subjects personally know any judge or former judge.

Empirical Claim 2 posited that candidates for judicial office who directly solicit campaign funds will diminish their own integrity and produce a lower sense of institutional trust and legitimacy than candidates who is raise campaign funds indirectly through their campaign committees. This was the primary empirical claim in Williams-Yulee and a major source of dispute between the majority and dissent. The claim is tested in Table 4 by comparing the outcome responses of subjects who were presented with vignettes in which Anderson personally solicited funds (direct solicitation) with subjects who were given vignettes in which Anderson's campaign committee was soliciting funds (indirect solicitation). Surprisingly, the estimated treatment effect of a judge directly asking for contributions on the impartiality and legitimacy measures is negative, although the size of the effect is small in both cases. On average, subjects believed that judges who have their campaign committees ask for money are less impartial and their court systems less legitimate than judges who do so themselves. Although these estimated effects are not significant, even the upper range of estimates within the 95% confidence interval suggest that the effect is negligible at best. Direct solicitation apparently does increase the perception of bias resulting from donations, but the effect is also not significant.

Table 4a: Indirect Solicitation v. Direct Solicitation

Outcome	Mean of Indirect Solicitation (n = 382)	Mean of Direct Solicitation (n = 373)	Estimated ATE (Regression w/ Covariates**)	p-value (2-tailed)	Lower 95% Confidence Interval	Upper 95% Confidence Interval
Fair and Impartial Judge	2.008	1.986	-0.029	0.652	-0.094	0.035
Legitimate Court	2.136	2.044	-0.083	0.218	-0.152	-0.016
Bias Due To Donation	2.727	2.825	0.100	0.107	0.038	0.162

* When accounting for the ten separate tests run on each of the three outcome measurements, the Šidák correction results in an alpha level of .005. **Coefficients and p-values in bold fall below the Šidák-corrected significance level.**

** Pre-treatment covariates include: age, sex, education, income, race (white), political ideology, political party affiliation (Democrat), previous presidential vote (Trump), whether the subjects have been personally involved in a court case, and whether the subjects personally know any judge or former judge.

Empirical Claim 3 posited that candidates for judicial office who send thank you notes to campaign donors will not diminish their own integrity and produce a lower sense of institutional trust and legitimacy than candidates who do not send thank you letters. Specifically, if thank you letters are as detrimental to legitimacy and impartiality as direct solicitation, then the law in question in *Williams-Yulee* would be under-inclusive. This claim is tested by comparing the outcome responses of subjects who were presented with vignettes that had Anderson sending thank you letters, with subjects who were given vignettes in which Anderson's did not send thank you letters. As with the results for empirical claim 2, the presence of purportedly detrimental campaign activities—thank you notes in this case—result in slightly higher perceptions of legitimacy and impartiality. This suggests that if there is an effect of thank you letters, it is positive, at least in regard to judicial perception. However, the substantive size of the effects is negligible and not statistically significant, as is the negative estimated effect on bias.

Table 5a: No Thank You Letter v. Thank You Letter

Outcome	Mean of No Letter (n = 386)	Mean of Thank You Letter (n = 369)	Estimated ATE (Regression w/ Covariates**)	p-value (2-tailed)	Lower 95% Confidence Interval	Upper 95% Confidence Interval
Fair and Impartial Judge	2.045	1.948	-0.090	0.167	-0.155	-0.025
Legitimate Court	2.111	2.069	-0.055	0.423	-0.123	0.014
Bias Due To Donation	2.815	2.734	-0.060	0.333	-0.122	0.002

* When accounting for the ten separate tests run on each of the three outcome measurements, the Šidák correction results in an alpha level of .005. **Coefficients and p-values in bold fall below the Šidák-corrected significance level.**

** Pre-treatment covariates include: age, sex, education, income, race (white), political ideology, political party affiliation (Democrat), previous presidential vote (Trump), whether the subjects have been personally involved in a court case, and whether the subjects personally know any judge or former judge.

In addition to the impact that judicial campaigns have on public perceptions of the judiciary, the majority in in *Williams-Yulee* was also convinced by Florida's argument that direct solicitation likely puts an undue burden on attorneys to donate to judges, for fear of being disadvantaged in the courtroom. Specifically, Empirical Claim 4 posits that attorneys are more pressured and therefore more likely to "donate" to a judicial candidate when that candidate directly solicits campaign funds than if that candidate solicited campaign funds indirectly through his campaign committee.

Of the four claims discussed in this Chapter, this claim is the least amenable to survey experiments, particularly when the survey respondents are not attorneys. While the primary outcome for the other hypotheses are individual perceptions, which can be captured in a survey, accurately testing the effect of campaign activities on the amount of "pressure" that attorneys feel, would ideally involve measuring the donation behavior of actual attorneys. However, to provide preliminary information on this relationship, the subjects' willingness to donate to Anderson's campaign was measured when he, as opposed to his committee, directly asked for

donations. As shown in Table 6a, the difference in the willingness to donate between the two conditions is negligible (.03 points on a 7-point scale) and not statistically significant.

Table 6a: v. Non-Solicitation v. Solicitation (Likelihood of Donating)

Outcome	Mean of Non-Solicitation (n = 382)	Mean of Solicitation (n = 373)	Estimated ATE (Regression w/ Covariates*)	p-value (2-tailed)	Lower 95% Confidence Interval	Upper 95% Confidence Interval
Likelihood of Donating (7-point scale)	3.739	3.702	-0.026	0.860	-0.173	0.121

* Pre-treatment covariates include: age, sex, education, income, race (white), state of residence, political ideology, political party affiliation (Democrat), previous presidential vote (Trump), whether the subjects have been personally involved in a court case, and whether the subjects personally know any judge or former judge.

D. Additional Hypotheses:

As described above, the survey vignette also included experimental elements relating to judicial elections and campaign finance that were not explicitly discussed in the *Williams-Yulee* opinions. These elements allow for tests of additional empirical hypotheses that build off of previous research and lay a groundwork for future studies on judicial elections. As with the empirical claims tested above, estimated treatment effects are calculated for the three primary outcomes of interest using multivariate regression analysis that controls for the pre-treatment covariates collected at the beginning of the survey.

Additional Hypothesis 1: Subjects presented with a judge or campaign committee who is actively raising campaign funds from attorneys will have a lower sense of institutional trust and legitimacy than subjects presented with a judge or campaign committee who is actively raising campaign funds from citizens.

Additional Hypothesis 2: Subjects presented with a judge or campaign committee who solicits campaign funds face-to-face (at a fundraising event) will have a lower sense of institutional trust and legitimacy than subjects presented with a judge or campaign committee who solicits campaign funds via letters.

The first two additional hypotheses build off of the issues that the justices were concerned with in *Williams-Yulee*. If, as the majority posits, direct solicitation is dangerous in

part because of how it will pressure attorneys to donate to judges, the public may perceive judges who target attorneys as being less impartial or unbiased than those who target citizens.

Table 7a: Solicitation of Attorneys v. Solicitation of Citizens

Outcome	Mean of Non-Solicitation of Citizens (n = 377)	Mean of Solicitation of Attorneys (n = 378)	Estimated ATE (Regression w/ Covariates**)	p-value (2-tailed)	Lower 95% Confidence Interval	Upper 95% Confidence Interval
Fair and Impartial Judge	1.965	2.030	0.080	0.220	0.015	0.144
Legitimate Court	2.038	2.143	0.106	0.120	0.038	0.174
Bias Due To Donation	2.778	2.774	0.011	0.857	-0.051	0.073

* When accounting for the ten separate tests run on each of the three outcome measurements, the Šidák correction results in an alpha level of .005. **Coefficients and p-values in bold fall below the Šidák-corrected significance level.**

** Pre-treatment covariates include: age, sex, education, income, race (white), political ideology, political party affiliation (Democrat), previous presidential vote (Trump), whether the subjects have been personally involved in a court case, and whether the subjects personally know any judge or former judge.

Table 8a: Letter Solicitation v. Face-to-Face Solicitation

Outcome	Mean of Letter Solicitation (n = 378)	Mean of Face-to-Face Solicitation (n = 377)	Estimated ATE (Regression w/ Covariates**)	p-value (2-tailed)	Lower 95% Confidence Interval	Upper 95% Confidence Interval
Fair and Impartial Judge	2.008	1.987	-0.063	0.337	-0.003	0.128
Legitimate Court	2.092	2.088	-0.036	0.605	-0.033	0.104
Bias Due To Donation	2.772	2.779	0.027	0.670	-0.089	0.036

* When accounting for the ten separate tests run on each of the three outcome measurements, the Šidák correction results in an alpha level of .005. **Coefficients and p-values in bold fall below the Šidák-corrected significance level.**

** Pre-treatment covariates include: age, sex, education, income, race (white), political ideology, political party affiliation (Democrat), previous presidential vote (Trump), whether the subjects have been personally involved in a court case, and whether the subjects personally know any judge or former judge.

Similarly, asking for donations in person may be seen as a more direct form of solicitation than asking for donations via campaign letters and may therefore influence how the public views judges who make such appeals. Tables 7a and 8a test these hypotheses. Given the results of the tests on Empirical Claim 2, it is perhaps not surprising that there are no significant treatment effects on any of the outcomes.

Additional Hypothesis 3: Subjects who are told that the judge publicly announces a policy to recuse himself from any case in which one or more of the parties or attorneys had previously donated to his political campaign will have a higher sense of institutional trust and legitimacy than subjects who are told that the judge declined to announce such a policy.

Previous survey experiments by Gibson and Caldeira have shown that while public perceptions of judicial legitimacy are damaged when judges hear cases in which a campaign donor is a participant, much (but not all) of that legitimacy can be recovered if the judge recuses from such cases.[71] Consistent with these findings, Table 9a indicates that subjects perceived the judges who refuse to recuse from cases that feature a donor attorney as moderately more impartial (.43 points) than judges who made a policy of recusing from such cases. Similarly, the court to which the judge was elected is seen as more legitimate (.32 points), and the judge is believed to be less biased (.21 points) when the judge recuses from cases. All of these results are significant at levels that survive a multiple comparison correction.

To determine whether this effect is driven by a proactive policy to recuse or the refusal to recuse, a vignette variation in which no mention of the judge's recusal policy is included. The regression results in Table 10a compare the perceptions of subjects who were not told a judge's recusal policy against those who were told that the judge has publicly stated that he will recuse. None of the outcomes are substantively or statistically significant, suggesting that the effects

[71] James L. Gibson & Gregory A. Caldeira, *Judicial Impartiality, Campaign Contributions, and Recusals: Results from a National Survey*, 10 J. EMP. L. STUD. 76 (2013); James L. Gibson & Gregory A. Caldeira, *Campaign Support, Conflicts of Interest, and Judicial Impartiality: Can Recusals Rescue the Legitimacy of Courts?,* 74 J. POL. 18 (2012).

presented in Table 9 are likely a result of both an aversion to a judge declining to recuse and an appreciation of a judge who positively states that he will recuse.

Table 9a: Policy of Non-Recusal v. Policy of Recusal

Outcome	Mean of Non-Recusal (n = 247)	Mean of Recusal (n = 265)	Estimated ATE (Regression w/ Covariates**)	p-value (2-tailed)	Lower 95% Confidence Interval	Upper 95% Confidence Interval
Fair and Impartial Judge	2.255	1.846	**-0.431**	**0.000**	-0.313	-0.349
Legitimate Court	2.321	2.032	**-0.329**	**0.000**	-0.415	-0.243
Bias Due To Donation	2.916	2.692	**-0.209**	**0.006**	-0.285	-0.133

* When accounting for the ten separate tests run on each of the three outcome measurements, the Šidák correction results in an alpha level of .005. **Coefficients and p-values in bold fall below the Šidák-corrected significance level.**
** Pre-treatment covariates include: age, sex, education, income, race (white), political ideology, political party affiliation (Democrat), previous presidential vote (Trump), whether the subjects have been personally involved in a court case, and whether the subjects personally know any judge or former judge.

Table 10a: No Mention of Recusal v. Policy of Recusal

Outcome	Mean of No Mention (n = 243)	Mean of Recusal (n = 265)	Estimated ATE (Regression w/ Covariates**)	p-value (2-tailed)	Lower 95% Confidence Interval	Upper 95% Confidence Interval
Fair and Impartial Judge	1.900	1.846	-0.037	0.613	-0.109	0.036
Legitimate Court	1.934	2.023	0.102	0.193	0.024	0.180
Bias Due To Donation	2.723	2.692	-0.062	0.422	-0.140	0.015

* When accounting for the ten separate tests run on each of the three outcome measurements, the Šidák correction results in an alpha level of .005. **Coefficients and p-values in bold fall below the Šidák-corrected significance level.**
** Pre-treatment covariates include: age, sex, education, income, race (white), political ideology, political party affiliation (Democrat), previous presidential vote (Trump), whether the subjects have been personally involved in a court case, and whether the subjects personally know any judge or former judge.

Additional Hypothesis 4: Subjects who are told that the judge's fundraising efforts led to large average donations or a large total donation amount will have a lower sense of

institutional trust and legitimacy than subjects who are told that the efforts led to small average donations or a small total donation amount.

The amount of fundraising and spending in judicial elections has drastically increased over the last thirty years,[72] and non-experimental empirical research has identified a strong, positive correlation between whether an individual (a party or an attorney) donates and how favorably the done judge rules on their case.[73] Similarly, national survey have shown that the majority of the broader public and nearly a majority of state-court judges themselves believe that money is influencing judicial behavior,[74] so it is likely that a judge who raises more money in his campaign will be seen as less legitimate, fair, and unbiased.

Table 11a compares the responses of subjects who were told that the judge raised a large amount of money ($101,000) or had large average donation amounts ($1,010 per donor) against those who were told that the judge raised a small amount of money ($4,500 total) or had small average donations ($45 per donor). Interestingly, the results of the multivariate regression suggest that increased money does not affect general perceptions of legitimacy or impartiality but does slightly (0.16 points) increase the likelihood that the subject believes the judge would

---

[72] *See, e.g.,* Chris W. Bonneau, *What Price Justice(s)? Understanding Campaign Spending in State Supreme Court Elections*, STATE POL. & POL'Y Q. 5: 107-125 (2005); Chris W. Bonneau, *The Dynamics of Campaign Spending in State Supreme Court Elections*, in RUNNING FOR JUDGE: THE RISING POLITICAL, FINANCIAL, AND LEGAL STAKES OF JUDICIAL ELECTIONS 63 fig.4.1 (Matthew J. Streb ed., 2007); Phillip L. Dubois, *Voting Cues in Nonpartisan Trial Court Elections: A Multivariate Assessment*, 18 L. & SOC. REV. 395 (1984) (reporting that the cost of election in the California Superior Courts (the trial courts) doubled from 1978 to 1982); California Commission on Campaign Financing, *The Price of Justice: A Los Angeles Area Case Study in Judicial Campaign Financing* (1995) (reporting that the "spending for Los Angeles County Superior Court races increased 22-fold" (at 51) between 1976 and 1994).

[73] *See, e.g.,* Morgan L. W. Hazelton, Jacob M. Montgomery, & Brendan Nyhan, *Does Public Financing Affect Judicial Behavior? Evidence from the North Carolina Supreme Court*, 44 AM. POL. RES. 587 (2015); Michael Heise, *The Complicated Business of State Supreme Court Elections: An Empirical Perspective*, Cornell Legal Studies Research Paper (2018).

[74] Justice at Stake, *Justice at Stake Frequency Questionnaire,* available at http://www.justiceatstake.org/media/cms/ JASNationalSurveyResults_6F537F99272D4.pdf (reporting that 76 percent of voters believe that campaign contributions have at least "some influence" on judicial behavior); Justice at Stake, *State Judges Frequency Questionnaire* (2002), available at http://www.justiceatstake.org /media/cms/JASJudgesSurveyResults_EA8838C0504A5.pdf. (reporting that 46 percent of state court judges stated that they believe campaign contributions influence judicial behavior).

be biased if a donor appeared in his courtroom. This finding is statistically significant at the .05 level but does not survive the multiple comparisons correction.

Table 11a: High Fundraising v. Low Fundraising

Outcome	Mean of Low Fundraising Amount (n = 250)	Mean of High Fundraising Amount (n = 241)	Estimated ATE (Regression w/ Covariates**)	p-value (2-tailed)	Lower 95% Confidence Interval	Upper 95% Confidence Interval
Fair and Impartial Judge	2.004	2.079	0.081	0.393	-0.002	0.163
Legitimate Court	2.065	2.134	0.072	0.329	-0.012	0.157
Bias Due To Donation	2.700	2.864	0.163	0.037	0.085	0.241

* When accounting for the ten separate tests run on each of the three outcome measurements, the Šidák correction results in an alpha level of .005. **Coefficients and p-values in bold fall below the Šidák-corrected significance level.**
** Pre-treatment covariates include: age, sex, education, income, race (white), political ideology, political party affiliation (Democrat), previous presidential vote (Trump), whether the subjects have been personally involved in a court case, and whether the subjects personally know any judge or former judge.

E. Florida Sub-group Analysis

Sub-group analyses for the 330 subjects residing in Florida allows for the tests that most closely match the fact pattern in *Williams-Yulee*. The results of the Florida sub-group analyses are presented in Tables 1b-6b (*see* Appendix B: Florida Sub-group Analysis). The outcomes on the four empirical claims are similar to those for the full sample—perceptions of the judiciary were not impacted by direct solicitation[75] or thank you letters, and the likelihood of donation was the same regardless of whether the judge asked for donations personally or did so through his campaign committee. The substantive size of the estimated ATEs were generally bigger for the Florida sample, but these effects were not statistically distinguishable from zero, possibly due to the smaller sample size and thus lower power of the tests.

[75] Direct solicitation did result in a .214-point increase in the respondent that believe a judge would be biased, and this estimated effect was significant at the .05 level. However, after accounting for multiple comparisons, this result is also indistinguishable from zero.

As with the analysis on the full sample, the Florida sub-sample test did identify some treatment effects pertaining to the additional hypotheses. Florida respondents were more skeptical of judges who asked for donations from attorneys than of those who asked for donations from citizens, although the neither of these effects survived the multiple comparison correction. Florida respondents were significantly more likely to think a judge would be biased in a case featuring a donor attorney (.345 points) when the average or total amount raised by the judge is high, and they were more skeptical across the board regarding judges who explicitly promised quid pro quo benefits to donors.

F. Limitations and Concerns:

i. Design-based Limitations: The key empirical questions in *Williams-Yulee* relate to the causal effect that various judicial campaign strategies have on the public's perception of the judiciary. The ideal empirical design for addressing these questions would satisfy the following three conditions: 1) random assignment, 2) treatment conditions that reproduce strategies used by judges in election campaigns, and 3) reliable measures for public perception of impartiality, legitimacy, and bias.

As an experiment, the survey featured in this Chapter satisfies the first of these requirements through the factorial randomization of vignette content. As a nationally representative survey, it also arguably satisfies the third condition. Although the majority in *Williams-Yulee* argued that "[t]he concept of public confidence in judicial integrity does not easily reduce to precise definition, nor does it lend itself to proof by documentary record,"[76] social scientists have long understood opinion surveys as a reliable method for collecting data on the public's trust in political institutions and beliefs regarding the impartiality of judges.[77] Additionally, the format and content of the questions included in this survey were adopted from

[76] Williams-Yulee, 135 S. Ct. at 1667.

[77] *See* the works of Gibson and Caldeira, *supra* note 39.

leading research in the area of judicial legitimacy and reflect the discipline's most tried and true approach for capturing such concepts.[78]

The question of how well a survey experiment can satisfy the second condition listed above is less clear. Both the majority and dissents in *Williams-Yulee* are concerned with the effect of *actual* campaign strategies used by *actual* judges. While the vignette used in this study is based largely on the campaign materials used by Williams-Yulee during her run for judgeship, the subjects are exposed to these materials in a survey context and are explicitly told that they are describing a hypothetical judicial candidate. Without a field experiment, it is unclear if subjects would react to the experimental treatments differently than they would in a natural setting. Additionally, citizens might be unaware of the campaign strategies of judges running for local seats in actual elections, in which case the effect that strategies such as direct fundraising will have on the public consciousness would be negligible, not because the public does not care, but because they are uninformed.

ii. The Weak-Null Problem: The empirical questions in *Williams-Yulee* are interesting because a null finding may be as interesting and valuable as an identified effect, and perhaps more so given that the majority assumed an effect. Most of the tests presented above resulted in estimated treatment effects that were statistically indistinguishable from zero. Finding a "true null" effect (a statistically significant but substantively small effect) for these treatments would provide more confidence that there actually no difference between the campaign strategies used in the vignettes. While the fact that the coefficients were small and the standard errors and the confidence intervals were tight, the "weak" nulls provides some confidence of negligible treatment effects, but to more reliable determine that there is no treatment effect, a larger and more powerful study is necessary.

[78] *See Id* and *supra* note 63.

V. Conclusions and Implications

The majority's decision in *Williams-Yulee* relied on the empirical conclusion that there is a substantial distinction between how the public views judges who raise campaign funds through direct solicitation and those who do so through indirect solicitation. The results of this study do not support this conclusion and therefore call into question the constitutional precedent set by *Williams-Yulee* and the subsequent decisions made by the U.S. Circuit and District Courts. While the evidence produced by this study should be understood with the limitations inherent in survey experiments in mind, this methodology is particularly well-suited for the questions raised in *Williams-Yulee* because surveys are best used to test public opinion and perception, which are the very outcomes that form the justification for the laws restricting direct solicitation by judicial candidates.

More broadly, the results of this study highlights value of critically evaluating the Supreme Court when it relies on empirical assumptions that are not supported by empirical data. While the Court cannot constrain its decisions purely to those questions that can and have been addressed by empiricists, continuing to rely on inherently empirical assumptions without the proper support will likely lead to flawed jurisprudence and lower public support (the very problem the Court sought to alleviate in *Williams-Yulee)*. Survey experiments such as the one featured in this Chapter are particularly well suited to this task, particularly when the empirical questions relate to the opinions and observations of the public.

Appendix A

Appendices to Chapter 1

Table 4: Treatment Effect on Recusal Across Covariate Categories

Covariate Category	Sub-Category	Control Sub-Category Proportion	Treatment Sub-Category Proportion	Sub-Category Estimated ATE**	Category Covariate Difference
Case Type	Divorce	0/6 (0%)	1/6 (17%)	.17	.20
---	Non-Divorce	2/24 (8%)	1/22 (5%)	-.03	
Judge Gender	Female	1/7 (14%)	1/9 (11%)	-.03	.04
---	Male	1/23 (4%)	1/19 (5%)	.01	
Judge Gender v. Attorney Gender	Same	1/19 (5%)	1/18 (6%)	.01	.00
---	Not Same	1/11 (9%)	1/10 (10%)	.01	
Previous Election Opposition	Unopposed	0/11 (0%)	0/12 (0%)	.00	.02
---	Opposed	2/19 (11%)	2/16 (13%)	.02	
Donation Amount	< $350	2/21 (10%)	2/18 (11%)	.01	.01
---	> $350	0/9 (0%)	0/10 (0%)	.00	
Donation Proportion (of Total External Donations*)	< 5%	1/20 (5%)	1/20 (5%)	.00	.03
---	> 5%	1/10 (10%)	1/8 (13%)	.03	

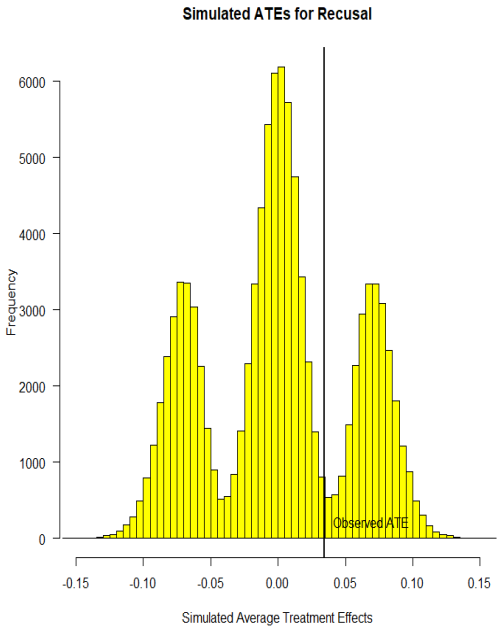
* External Donations exclude any donations made by the candidate or individuals with the same surname as the candidate.
** This estimate of the ATE is calculated without weighting or the use of pre-treatment covariates.

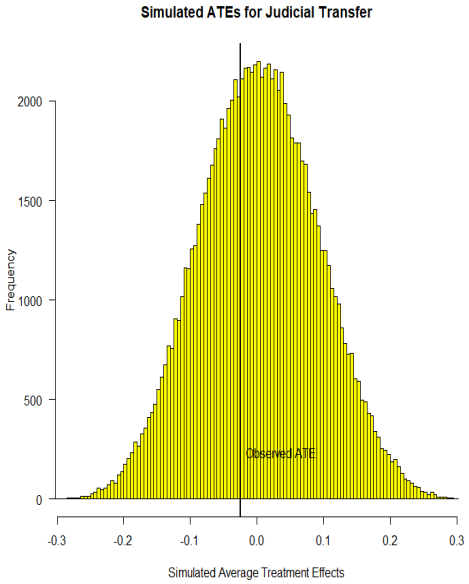
Table 5: Treatment Effect on Disclosure Across Covariate Categories

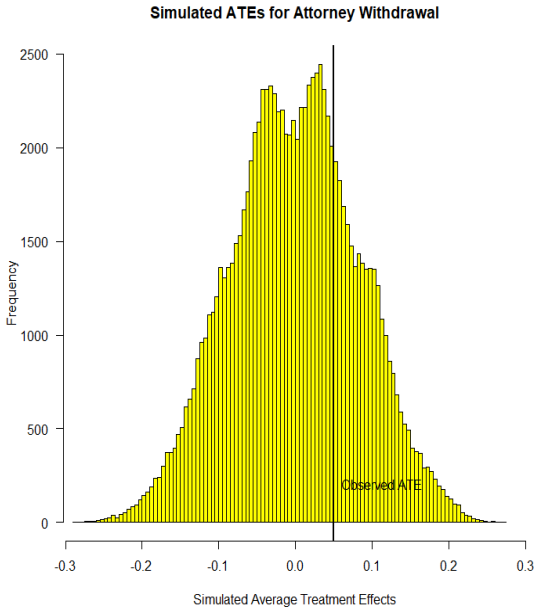
Covariate Category	Sub-Category	Control Sub-Category Proportion	Treatment Sub-Category Proportion	Sub-Category Estimated ATE**	Category Covariate Difference
Case Type	Divorce	0/6 (0%)	4/6 (66%)	.66	.43
---	Non-Divorce	0/24 (0%)	5/22 (23%)	.23	
Judge Gender	Female	0/7 (0%)	2/9 (22%)	.22	.15
---	Male	0/23 (0%)	7/19 (37%)	.37	
Judge Gender v. Att. Gender	Same	0/19 (0%)	4/18 (22%)	.22	.28
---	Not Same	0/11 (0%)	5/10 (50%)	.50	
Previous Election Opposition	Unopposed	0/11 (0%)	5/12 (42%)	.42	.17
---	Opposed	0/19 (0%)	4/16 (25%)	.25	
Donation Amount	< $350	0/21 (0%)	7/18 (39%)	.39	.19
---	> $350	0/9 (0%)	2/10 (20%)	.20	
Donation Proportion (of Total External Donations*)	< 5%	0/20 (0%)	7/20 (35%)	.35	.10
---	> 5%	0/10 (0%)	2/8 (25%)	.25	

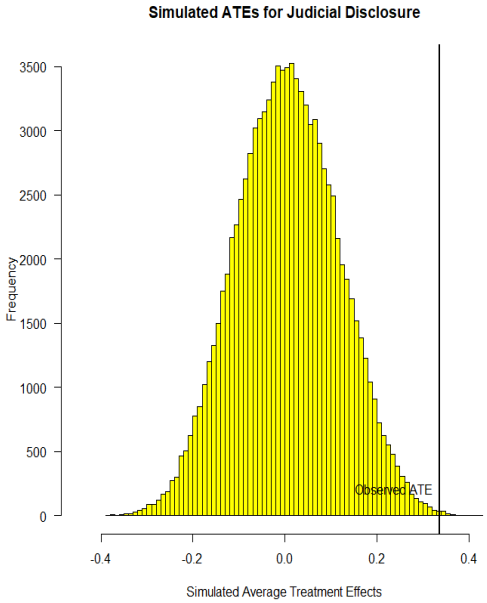
* External Donations exclude any donations made by the candidate or individuals with the same surname as the candidate.
** This estimate of the ATE is calculated without weighting or the use of pre-treatment covariates.

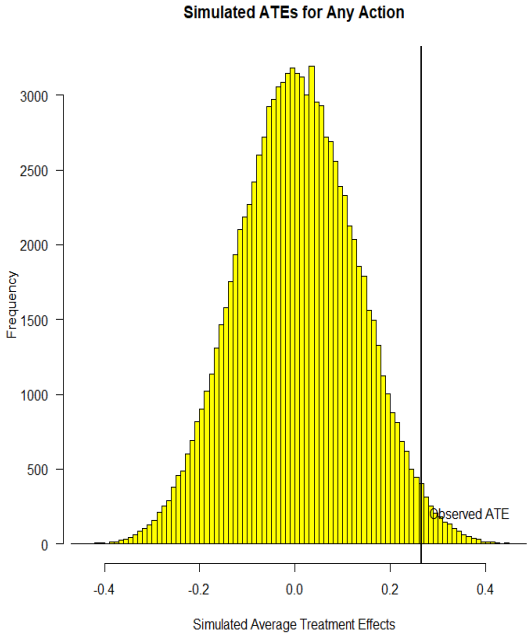
Figure 5: Plots of Simulated ATEs (Wisconsin) [10,000 Permutations]

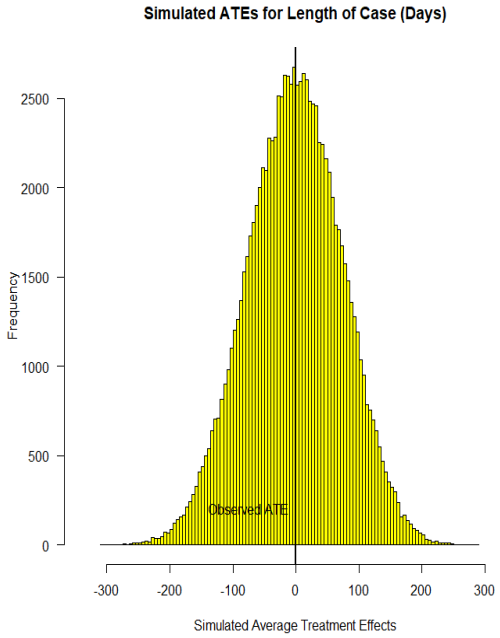









Simulated ATEs for Any Action
Frequency
3000
2500
2000
1500
1000
500
0
Observed ATE
-0.4
-0.2
0.0
0.2
0.4
Simulated Average Treatment Effects


Simulated ATEs for Length of Case (Days)
Frequency
2500
2000
1500
1000
500
0
Observed ATE
-300
-200
-100
0
100
200
300
Simulated Average Treatment Effects

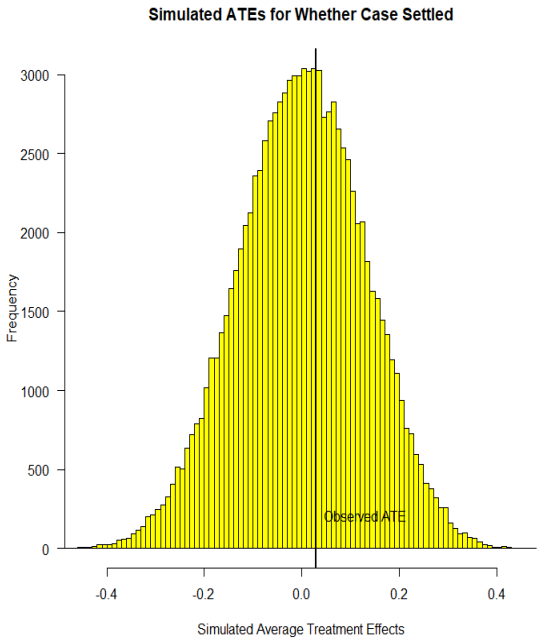

Simulated ATEs for Whether Case Settled
Frequency
3000
2500
2000
1500
1000
500
0
Observed ATE
-0.4
-0.2
0.0
0.2
0.4
Simulated Average Treatment Effects

Figure 6: Harris County Attorney Verification Protocol Script

1. Does the Texas Bar database have an individual with the same first and last name as the individual in the case and the individual in the campaign finance data (note: there was never any discrepancy between the names in the cases and the names in the bar database)?
    - If Yes: Move on to next step.
    - If No: Does the Texas Bar database have an individual with the same last name and a similar first name as the individual in the case and the individual in the campaign finance data? (Examples: Bob and Robert, Rusty and Russel, Tony and Anthony, and anyone with a middle name that matches a first name.)
        - If Yes: Move onto next step.
        - If No: Exclude from the sample.
2. Is there more than one individual in the Texas Bar database that has that first and last name or a similar name? (Examples: There are two William Stouts, two Keith Fletchers, and six Jose Lopez on the Texas Bar database.)
    - If No: Include in the sample as a treatable case.
    - If Yes: Move onto next step.
3. Do any of the individuals of this name on the Texas Bar database list the same law firm as the individual on the campaign finance list?
    - If Yes: Move onto step 4.
    - If No: Move onto step 5.
4. Is there more than one individual of this name on the Texas Bar database with the same law firm as the individual on the campaign finance list? (Examples: There are two Keith Fletchers who work at the same firm--clearly father and son)
    - If Yes: Move onto step 5.
    - If No: Include in the sample as a treatable case.
5. Do any of the individuals of this name on the Texas Bar database list the same postal code as the individual on the campaign finance list?
    - If Yes: Move onto step 6.
    - If No: Move onto step 7.
6. Is there more than one individual of this name on the Texas Bar database with the same postal code as the individual on the campaign finance list?
    - If Yes: Move onto step 6.
    - If No: Include in the sample as a treatable case.
7. Is there any good evidence?
    - Examples of good evidence: After going to the attorney's firm page (the firm that is listed in the Bar database), we see that she worked at the firm listed in the campaign finance data at the time she made the donation. OR The firm names are not an exact match, but are clearly the same firm (Thorley and partners = Thorley LLC) OR There are two attorneys of the same name working at the same firm, but the donor attorney is listed as a partner in the campaign finance data and only one of these two individuals is a partner. OR There are two attorneys of the same name working at the same firm, but only one was an attorney at the time the donation was made.
    - No: Exclude from the sample.
    - Yes: Include in the sample as a treatable case and make note of this evidence in the excel file.

Figure 7: Sample of Assignment Blocks (Wisconsin)

Blocks	Number of Cases in Treatment	Number of Cases in Control	Total Number of Cases
Block 1 (1 judge, 3	1	2	3
Block 2 (1 judge, 2	1	1	2
Block 3 (1 judge, 2	1	1	2
Block 4 (1 judge, 2	1	1	2
Block 5 (1 judge, 2	1	1	2
Block 6 (5 judges, 5	2	3	5
Block 7 (2 judges, 2	1	1	2
Block 8 (2 judges, 2	1	1	2
Block 9 (1 judge, 1 case)	0	1	1
Block 10 (1 judge, 2	1	1	2
Block 11 (1 judge, 1	0	1	1
Block 12 (1 judge, 1	0	1	1
Block 13 (1 judge, 1	1	0	1
Block 14 (3 judges, 3	2	1	3
Block 15 (2 judges, 2	1	1	2
Block 16 (1 judge, 1	1	0	1
Block 17 (1 judge, 1	0	1	1
Block 18 (1 judge, 1	0	1	1
Block 19 (1 judge, 1	0	1	1
Block 20 (1 judge, 1	0	1	1
Block 21 (1 judge, 1	1	0	1
Block 22 (1 judge, 1	0	1	1
Block 23 (3 judges, 3	2	1	3
Block 24 (1 judge, 1	0	1	1
Block 25 (2 judges, 2	0	1	1
Block 26 (1 judge, 1	0	1	1
Block 27 (1 judge, 1	1	0	1
Block 28 (1 judge, 1	1	0	1

Appendix B

Appendices to Chapter 3

Text of Original Williams-Yulee Letter[1]

Reply to Tampa Post Office Box 340031
Tampa, Florida 33694
email: lwilliams94@verizon.net
www.vote4vulee.com

"Bringing Diversity to the Judicial Bench"

RE: Elect Lanell Williams-Yulee For County Court
Judge Group 10 and Campaign Fundraiser

Dear Friend:

I have served as a public servant for this community as Public Defender as well as a Prosecutor for the past 18 years. Having been involved in various civic activities such as "The Great American Teach In," Inns Of Court, Pro Bono Attorney, Metropolitan Ministries outreach program, as well as a mentor for various young men and women residing within Hillsborough County, I have long worked for positive change in Tampa. With the support of my family, I now feel that the time has come for me to seek elected office. I want to bring fresh ideas and positive solutions to the Judicial bench. I am certain that I can uphold the Laws, Statutes, Ordinances as prescribed by the Constitution Of the State Of Florida as well as the Constitution of the United States Of America. I am confident that I can serve as a positive attribute to the Thirteenth Judicial Circuit by running for County Court Judge, Group 10. To succeed in this effort, I need to mount an aggressive campaign. I'm inviting the people that know me best to join my campaign and help make a real difference. An early contribution of $25, $50, $100, $250, or $500, made payable to "Lanell Williams-Yulee Campaign for County Judge", will help raise the initial funds needed to launch the campaign and get our message out to the public. I ask for your support In meeting the primary election fund raiser goals. Thank you in advance for your support.

Sincerely,
/s/
Lanell Williams-Yulee, Esq.

Political Advertisement paid for and approved by
Lanell Williams-Yulee, Nonpartisan, for County
Judge, Group 10

[1] Petition for a Writ of Certiorari at 31a–32a, Williams-Yulee v. The Florida Bar, 138 So.3d 379 (No. 13-1499) (June 17, 2014), 2014 WL 2769040.

Florida Sub-Group Analysis

Table 1b: Sample Demographics (Florida)

Respondent Characteristic	Mean/ Proportion	Median	Minimum	Maximum
Age (Years)	44.79	44	18	92
Sex (Female)	0.53	Female	NA	NA
Education (College or Higher)	0.42	Some College	8th Grade	Post-graduate Degree
Income (US Dollars)	$55,877	$45,000	$0	$275,000
Race (White)	0.74	White	NA	NA
Ideology (5-point Scale; Low = Conservative)	2.91	3	1	5
Political Party Affiliation (Democrat)	0.35	Republican	NA	NA
Vote in Last Presidential Election (Trump)	0.35	Trump	NA	NA
Personal Experience in Court (Yes)	0.29	No Experience	NA	NA
Personally Know a Judge (Yes)	0.12	Do Not Know Judge	NA	NA

Table 2b: Quid Pro Quo v. Regular Solicitation (Florida)

Outcome	Mean of Regular Solicitation (n = 303)	Mean of Quid Pro Quo (n = 27)	Estimated ATE (Regression w/ Covariates**)	P-value* (2-tailed)	Lower 95% Confidence Interval	Upper 95% Confidence Interval
Fair and Impartial Judge	1.950	2.85	**0.941**	**0.000**	0.755	1.127
Legitimate Court	2.127	3.000	**0.941**	**0.000**	0.74	1.142
Bias Due To Donation	2.709	3.370	**0.545**	**0.002**	0.370	0.720
Likelihood of Donating	3.586	3.950	0.340	0.208	0.071	0.610

* When accounting for the ten separate tests run on each of the three outcome measurements, the Šidák correction results in an alpha level of .005. **Coefficients and p-values in bold fall below the Šidák-corrected significance level.**

** Pre-treatment covariates include: age, sex, education, income, race, state of residence, political ideology, political party affiliation, previous presidential vote, whether the subjects have been personally involved in a court case, and whether the subjects personally know any judge or former judge.

Table 3b: Non-Solicitation v. Solicitation (Florida)

Outcome	Mean of Non-Solicitation (n = 64)	Mean of Solicitation (n = 239)	Estimated ATE (Regression w/ Covariates**)	p-value (2-tailed)	Lower 95% Confidence Interval	Upper 95% Confidence Interval
Fair and Impartial Judge	1.891	1.966	0.053	0.667	-0.070	0.175
Legitimate Court	2.105	2.157	0.130	0.332	-0.004	0.263
Bias Due To Donation	2.718	2.706	0.019	0.872	-0.098	0.136

* When accounting for the ten separate tests run on each of the three outcome measurements, the Šidák correction results in an alpha level of .005. **Coefficients and p-values in bold fall below the Šidák-corrected significance level.**

** Pre-treatment covariates include: age, sex, education, income, race, state of residence, political ideology, political party affiliation, previous presidential vote, whether the subjects have been personally involved in a court case, and whether the subjects personally know any judge or former judge.

Table 4b: Indirect Solicitation v. Direct Solicitation (Florida)

Outcome	Mean of Indirect Solicitation (n = 117)	Mean of Direct Solicitation (n = 122)	Estimated ATE (Regression w/ Covariates**)	p-value (2-tailed)	Lower 95% Confidence Interval	Upper 95% Confidence Interval
Fair and Impartial Judge	1.913	2.016	0.106	0.355	-0.008	0.221
Legitimate Court	2.129	2.184	0.126	0.314	0.001	0.250
Bias Due To Donation	2.600	2.808	0.215	0.043	0.109	0.320

* When accounting for the ten separate tests run on each of the three outcome measurements, the Šidák correction results in an alpha level of .005. **Coefficients and p-values in bold fall below the Šidák-corrected significance level.**

** Pre-treatment covariates include: age, sex, education, income, race (white), political ideology, political party affiliation (Democrat), previous presidential vote (Trump), whether the subjects have been personally involved in a court case, and whether the subjects personally know any judge or former judge.

Table 5b: No Thank You Letter v. Thank You Letter (Florida)

Outcome	Mean of No Letter (n = 123)	Mean of Thank You Letter (n = 116)	Estimated ATE (Regression w/ Covariates**)	p-value (2-tailed)	Lower 95% Confidence Interval	Upper 95% Confidence Interval
Fair and Impartial Judge	1.983	1.947	0.008	0.937	-0.102	0.120
Legitimate Court	2.131	2.186	0.072	0.554	-0.049	0.192
Bias Due To Donation	2.672	2.743	0.102	0.325	-0.001	0.206

* When accounting for the ten separate tests run on each of the three outcome measurements, the Šidák correction results in an alpha level of .005. **Coefficients and p-values in bold fall below the Šidák-corrected significance level.**

** Pre-treatment covariates include: age, sex, education, income, race (white), political ideology, political party affiliation (Democrat), previous presidential vote (Trump), whether the subjects have been personally involved in a court case, and whether the subjects personally know any judge or former judge.

Table 6b: Non-Solicitation v. Solicitation (Likelihood of Donating) (Florida)

Outcome	Mean of Non-Solicitation (n = 117)	Mean of Solicitation (n = 122)	Estimated ATE (Regression w/ Covariates**)	p-value (2-tailed)	Lower 95% Confidence Interval	Upper 95% Confidence Interval
Likelihood of Donating (7-point scale)	3.586	3.950	0.340	0.208	0.071	0.609

** Pre-treatment covariates include: age, sex, education, income, race (white), state of residence, political ideology, political party affiliation (Democrat), previous presidential vote (Trump), whether the subjects have been personally involved in a court case, and whether the subjects personally know any judge or former judge.

Survey Text and Logic

Note: Language in [brackets] did not appear to the survey subjects.

START OF SURVEY

[Start of Block 1: Consent]

[If subjects do not consent to taking the survey, they were not allowed to proceed.]

Thank you for choosing to take this survey!

In order for you to participate in this survey, we need you to read over and accept the consent form below.

When you are done reading the form, you can begin the survey by selecting the option, "I consent to take this survey." By selecting that option, you are indicating that the you have read the below information, have had the opportunity to have any questions about this study answered, are at least 18 years old, and agree to participate in this study.

Consent Form

Responsible Institution: Columbia University

Investigators: Dane Thorley and Donald Green (Principal Investigator) IRB Protocol # AAAR7367

Purpose: We are conducting a research study to examine people's opinions about individuals and their behavior.

Procedures: Participation in this study will involve answering some basic demographic questions, answering some questions about your political views, reading a short scenario, and answering some questions about what you read.

Risks and Benefits: Participants in this study may experience mild boredom due to participating in this study. Although this study will not benefit you personally, we hope that our results will add to the knowledge about legal disputes.

Confidentiality: All of your responses will be anonymous. We will not ask for, or record, your name or other information that could lead back to you. Only the researchers involved in this study and those responsible for research oversight will have access to the information you provide. When we publish any results from this study we will do so in a way that does not identify you.

Voluntary Participation: Participation in this study is completely voluntary. You are free to decline to participate, to end participation at any time for any reason, or to refuse to answer any individual question. Refusing to participate will involve no penalty or loss of benefits or compensation to which you are otherwise entitled or affect your relationship with any person or institution involved in the study.
Questions: If you have any questions about this study, you may contact Dane Thorley at drt2121@columbia.edu.

If you would like to talk with someone other than the researchers to discuss problems or concerns, to discuss situations in the event that a member of the research team is not available, or to discuss your rights as a research participant, you may contact the Columbia University Institutional Review Board (IRB) Administrator. They can be reached at (+1) 212-851-7040 or at: Columbia University Morningside IRB, 615 West 131st Street, 3rd Floor (Mail Code 8716) New York, NY 10027 (email address askirb@columbia.edu). Please mention protocol number # AAAR7367 when you write to the IRB.

- o I consent to take this survey.
- o No. I do not consent, and I do not want to take this survey.

[End of Block 1: Consent]

Page Break

[Start of Block 2: I – Intro]

To begin, we have a few questions about you.

[End of Block 2: I – Intro]

Page Break

[Start of Block 3: I - Demographic Questions]

Age How old are you?
_____

Page Break

What is your sex?
- o Female

- Male
- Other

Page Break

What is your current education level?
- Eight grade or less
- Some high school
- High school graduate
- Some college
- College graduate
- Post-college education
- Other

Page Break

What is your current household income (in dollars)?
_____

Page Break

What is your race or ethnicity?
- Asian
- Black
- Caucasian (White)
- Hispanic
- Latin
- Middle-Eastern
- Native American
- Multiracial
- Other

Page Break

In what U.S. state do you currently live?
▼ Alabama ... Wyoming

Page Break

Do you consider yourself ideologically conservative, ideologically liberal, or somewhere in the middle?
- Strongly conservative
- Somewhat conservative
- Moderate (the middle)
- Somewhat liberal
- Strongly liberal

Page Break

Do you consider yourself a Democrat, Republican, or an Independent?
- Democrat
- Republican
- Independent/Unaffiliated
- Other

Page Break

Who did you vote for in the last US presidential election?
- Hillary Clinton
- Donald Trump
- Gary Johnson
- Jill Stein
- Evan McMullen
- Darrell Castle
- Other
- I did not vote

Page Break

Have you ever been personally involved in a court case?
- Yes
- No

Page Break

Do you personally know any judges or former judges?
- Yes
- No

[End of Block 3: I - Demographic Questions]

Page Break

[Start of Block 4: II - Intro]

Now you will be presented with a hypothetical (fictional) description of a judicial candidate running for election. We will then ask you for your opinions about this candidate and his election campaign.

Please read the description of the candidate and questions carefully. At the end of this survey, we will ask you a question about the details of this individual's campaign to make sure you have read closely.

[End of Block 4: II - Intro]

Page Break

[Start of Block 5: II - Experimental Vignette Part 1/3]

[Each subject was randomly assigned to one of the following texts for part 1/3 of the total vignette. All conditions in this part were assigned with equal probability, with the exception of the treatment condition (A.), which was assigned at twice the probability of all other sections. Assignments of conditions in this part were independent of assignments in other vignette parts.]

[A. Control]

Last year, Fairview County (a fictional county in Florida) held elections for the judges in their trial courts. A citizen and attorney of Fairview County named George Anderson decided that he would like to run for one of the judicial positions. Like most of the judicial candidates running in the election, George Anderson needed to raise money for his campaign.

In order to do this, Anderson included his name on the official election registry. While he did not solicit or ask for any donations, he was willing to take donations from individuals who knew he was running for office and wanted to support his campaign. He also created a website for his campaign, which included the following statement:

Dear Friends:

I have served as a public servant for this community as public defender as well as a prosecutor for the past 18 years. Having been involved in various civic activities such as “The Great American Teach In”, Inns of Court, pro bono legal assistance, Metropolitan Ministries outreach program, as well as a mentor for various young men and women residing within Fairview, I have long worked for positive change in this county and in Florida broadly.

With the support of my family, I now feel that the time has come for me to seek elected office. I want to bring fresh ideas and positive solutions to the Fairview County judicial branch. I am certain that I can uphold the laws, statutes, and ordinances as prescribed by the Constitution of The State Of Florida as well as the Constitution of the United States of America.

Thank you in advance for your support!

George Anderson

[B. Judge Sends Letter to Attorneys]

Last year, Fairview County (a fictional county in Florida) held elections for the judges in their trial court. A citizen and attorney of Fairview County named George Anderson decided that he would like to run for one of the judicial positions. Like most of the judicial candidates running in the election, George Anderson needed to raise money for his campaign.

In order to do this, Anderson sent the following letter to all of the attorneys in Fairview County:

Dear Friend:

I have served as a public servant for this community as public defender as well as a prosecutor for the past 18 years. Having been involved in various civic activities such as "The Great American Teach In", Inns of Court, pro bono legal assistance, Metropolitan Ministries outreach program, as well as a mentor for various young men and women residing within Fairview, I have long worked for positive change in this county and in Florida broadly.

With the support of my family, I now feel that the time has come for me to seek elected office. I want to bring fresh ideas and positive solutions to the Fairview County judicial branch. I am certain that I can uphold the laws, statutes, and ordinances as prescribed by the Constitution of The State Of Florida as well as the Constitution of the United States of America.

To succeed in this effort, I need to mount an aggressive campaign. I'm inviting the people that know me best to join my campaign and help make a real difference. An early contribution of $25, $50, $100, $250, or $500, made payable to the "George Anderson Campaign for County Judge" will help raise the initial funds needed to launch the campaign and get out our message to the public.

You can make online donations at Anderson_for_Judge.com or mail in contributions to:

George Anderson Campaign for Fairview County Judge, Fairview PO Box 100101, Fairview, Florida 33456.

Thank you in advance for your support!

George Anderson

[C. Judge Sends Letter to Citizens]

Last year, Fairview County (a fictional county in Florida) held elections for the judges in their trial court. A citizen and attorney of Fairview County named George Anderson decided that he would like to run for one of the judicial positions. Like most of the judicial candidates running in the election, George Anderson needed to raise money for his campaign.

In order to do this, Anderson sent the following letter to all of the citizens in Fairview County:

Dear Friend:

I have served as a public servant for this community as public defender as well as a prosecutor for the past 18 years. Having been involved in various civic activities such as "The Great American Teach In", Inns of Court, pro bono legal assistance, Metropolitan Ministries outreach program, as well as a mentor for various young men and women residing within Fairview, I have long worked for positive change in this county and in Florida broadly.

With the support of my family, I now feel that the time has come for me to seek elected office. I want to bring fresh ideas and positive solutions to the Fairview County judicial branch. I am certain that I can uphold the laws, statutes, and ordinances as prescribed by the Constitution of The State Of Florida as well as the Constitution of the United States of America.

To succeed in this effort, I need to mount an aggressive campaign. I'm inviting the people that know me best to join my campaign and help make a real difference. An early contribution of $25, $50, $100, $250, or $500, made payable to the "George Anderson Campaign for County Judge" will help raise the initial funds needed to launch the campaign and get out our message to the public.

You can make online donations at Anderson_for_Judge.com or mail in contributions to:

George Anderson Campaign for Fairview County Judge, Fairview PO Box 100101, Fairview, Florida 33456.

Thank you in advance for your support!

George Anderson

[D. Committee Sends Letter to Attorneys]

Last year, Fairview County (a fictional county in Florida) held elections for the judges in their trial court. A citizen and attorney of Fairview County named George Anderson decided that he would like to run for one of the judicial positions. Like most of the judicial candidates running in the election, George Anderson needed to raise money for his campaign.

In order to do this, Anderson had his campaign committee send the following letter to all of the attorneys in Fairview County:

Dear Friend:

George Anderson has served as a public servant for this community as public defender as well as a prosecutor for the past 18 years. Having been involved in various civic activities such as “The Great American Teach In”, Inns of Court, pro bono legal assistance, Metropolitan Ministries outreach program, as well as a mentor for various young men and women residing within Fairview, he has long worked for positive change in this county and in Florida broadly.

With the support of his family, George Anderson now feels that the time has come for him to seek elected office. He wants to bring fresh ideas and positive solutions to the Fairview County judicial branch. He is certain that he can uphold the laws, statutes, and ordinances as prescribed by the Constitution of The State Of Florida as well as the Constitution of the United States of America.

To succeed in this effort, he needs to mount an aggressive campaign. As his campaign committee, we are inviting the people that know him best to join his campaign and help make a real difference. An early contribution of $25, $50, $100, $250, or $500, made payable to “The George Anderson Campaign for County Judge” will help raise the initial funds needed to launch the campaign and get out our message to the public.

You can make online donations at Anderson_for_Judge.com or mail in contributions to:

The George Anderson Campaign for County Judge, Fairview PO Box 100101, Fairview, Florida 33456

Thank you in advance for your support!

The George Anderson Campaign for County Judge

[E. Committee Sends Letter to Citizens]

Last year, Fairview County (a fictional county in Florida) held elections for the judges in their trial court. A citizen and attorney of Fairview County named George Anderson decided that he would like to run for one of the judicial positions. Like most of the judicial candidates running in the election, George Anderson needed to raise money for his campaign.

In order to do this, Anderson had his campaign committee send the following letter to all of the citizens in Fairview County:

Dear Friend:

George Anderson has served as a public servant for this community as public defender as well as a prosecutor for the past 18 years. Having been involved in various civic activities such as “The Great American Teach In”, Inns of Court, pro bono legal assistance, Metropolitan Ministries outreach program, as well as a mentor for various young men and women residing within Fairview, he has long worked for positive change in this county and in Florida broadly.

With the support of his family, George Anderson now feels that the time has come for him to seek elected office. He wants to bring fresh ideas and positive solutions to the Fairview County judicial branch. He is certain that he can uphold the laws, statutes, and ordinances as prescribed

by the Constitution of The State Of Florida as well as the Constitution of the United States of America.

To succeed in this effort, he needs to mount an aggressive campaign. As his campaign committee, we are inviting the people that know him best to join his campaign and help make a real difference. An early contribution of $25, $50, $100, $250, or $500, made payable to "The George Anderson Campaign for County Judge" will help raise the initial funds needed to launch the campaign and get out our message to the public.

You can make online donations at Anderson_for_Judge.com or mail in contributions to:

The George Anderson Campaign for County Judge, Fairview PO Box 100101, Fairview, Florida 33456

Thank you in advance for your support! The George Anderson Campaign for County Judge

[F. Judge Asks Attorneys at Event]

Last year, Fairview County (a fictional county in Florida) held elections for the judges in their trial court. A citizen and attorney of Fairview County named George Anderson decided that he would like to run for one of the judicial positions. Like most of the judicial candidates running in the election, George Anderson needed to raise money for his campaign.

In order to do this, Anderson organized a campaign donation event, at which he personally asked guests (who were mostly attorneys in Fairview County) to donate to his campaign, generally saying something like this:

I have served as a public servant for this community as public defender as well as a prosecutor for the past 18 years. Having been involved in various civic activities such as "The Great American Teach In", Inns of Court, pro bono legal assistance, Metropolitan Ministries outreach program, as well as a mentor for various young men and women residing within Fairview, I have long worked for positive change in this county and in Florida broadly.

With the support of my family, I now feel that the time has come for me to seek elected office. I want to bring fresh ideas and positive solutions to the Fairview County judicial branch. I am certain that I can uphold the laws, statutes, and ordinances as prescribed by the Constitution of The State Of Florida as well as the Constitution of the United States of America.

To succeed in this effort, I need to mount an aggressive campaign. I'm inviting the people that know me best to join my campaign and help make a real difference. An early contribution of $25, $50, $100, $250, or $500, made payable to the "George Anderson Campaign for County Judge" will help raise the initial funds needed to launch the campaign and get out our message to the public.

You can make donations at the table over there or visit my website and make payments online.

Thank you in advance for your support!

[G. Judge Asks Citizens at Event]

Last year, Fairview County (a fictional county in Florida) held elections for the judges in their trial court. A citizen and attorney of Fairview County named George Anderson decided that he would like to run for one of the judicial positions. Like most of the judicial candidates running in the election, George Anderson needed to raise money for his campaign.

In order to do this, Anderson organized a campaign donation event, at which he personally asked guests (who were mostly citizens of Fairview County) to donate to his campaign, generally saying something like this:

I have served as a public servant for this community as public defender as well as a prosecutor for the past 18 years. Having been involved in various civic activities such as "The Great American Teach In", Inns of Court, pro bono legal assistance, Metropolitan Ministries outreach program, as well as a mentor for various young men and women residing within Fairview, I have long worked for positive change in this county and in Florida broadly.

With the support of my family, I now feel that the time has come for me to seek elected office. I want to bring fresh ideas and positive solutions to the Fairview County judicial branch. I am certain that I can uphold the laws, statutes, and ordinances as prescribed by the Constitution of The State Of Florida as well as the Constitution of the United States of America.

To succeed in this effort, I need to mount an aggressive campaign. I'm inviting the people that know me best to join my campaign and help make a real difference. An early contribution of $25, $50, $100, $250, or $500, made payable to the "George Anderson Campaign for County Judge" will help raise the initial funds needed to launch the campaign and get out our message to the public.

You can make donations at the table over there or visit my website and make payments online.

Thank you in advance for your support!

[H. Committee Asks Attorneys at Event]

Last year, Fairview County (a fictional county in Florida) held elections for the judges in their trial court. A citizen and attorney of Fairview County named George Anderson decided that he would like to run for one of the judicial positions. Like most of the judicial candidates running in the election, George Anderson needed to raise money for his campaign.

In order to do this, Anderson organized a campaign donation event (Anderson was not in attendance), at which his campaign committee asked guests (who were mostly attorneys in Fairview County) to donate to his campaign, generally saying something like this:

George Anderson has served as a public servant for this community as public defender as well as a prosecutor for the past 18 years. Having been involved in various civic activities such as "The Great American Teach In", Inns of Court, pro bono legal assistance, Metropolitan Ministries outreach program, as well as a mentor for various young men and women residing within Fairview, he has long worked for positive change in this county and in Florida broadly.

With the support of his family, George Anderson now feels that the time has come for him to seek elected office. He wants to bring fresh ideas and positive solutions to the Fairview County judicial branch. He is certain that he can uphold the laws, statutes, and ordinances as prescribed by the Constitution of The State Of Florida as well as the Constitution of the United States of America.

To succeed in this effort, he needs to mount an aggressive campaign. As his campaign committee, we are inviting the people that know him best to join his campaign and help make a real difference. An early contribution of $25, $50, $100, $250, or $500, made payable to "The George Anderson Campaign for County Judge" will help raise the initial funds needed to launch the campaign and get out our message to the public.

You can make donations at the table over there or visit his website and make payments online.

Thank you in advance for your support!

[I. Committee Asks Citizens at Event]

Last year, Fairview County (a fictional county in Florida) held elections for the judges in their trial court. A citizen and attorney of Fairview County named George Anderson decided that he would like to run for one of the judicial positions. Like most of the judicial candidates running in the election, George Anderson needed to raise money for his campaign.

In order to do this, Anderson organized a campaign donation event (Anderson was not in attendance), at which his campaign committee asked guests (who were mostly citizens of Fairview County) to donate to his campaign, generally saying something like this:

George Anderson has served as a public servant for this community as public defender as well as a prosecutor for the past 18 years. Having been involved in various civic activities such as "The Great American Teach In", Inns of Court, pro bono legal assistance, Metropolitan Ministries outreach program, as well as a mentor for various young men and women residing within Fairview, he has long worked for positive change in this county and in Florida broadly.

With the support of his family, George Anderson now feels that the time has come for him to seek elected office. He wants to bring fresh ideas and positive solutions to the Fairview County judicial branch. He is certain that he can uphold the laws, statutes, and ordinances as prescribed by the Constitution of The State Of Florida as well as the Constitution of the United States of America.

To succeed in this effort, he needs to mount an aggressive campaign. As his campaign committee, we are inviting the people that know him best to join his campaign and help make a real difference. An early contribution of $25, $50, $100, $250, or $500, made payable to "The George Anderson Campaign for County Judge" will help raise the initial funds needed to launch the campaign and get out our message to the public.

You can make donations at the table over there or visit his website and make payments online.

Thank you in advance for your support!

[J. Explicit Quid Pro Quo]

Last year, Fairview County (a fictional county in Florida) held elections for the judges in their trial court. A citizen and attorney of Fairview County named George Anderson decided that he would like to run for one of the judicial positions. Like most of the judicial candidates running in the election, George Anderson needed to raise money for his campaign.

In order to do this, Anderson directly asked a number of prominent attorneys in Fairview County for campaign money, promising to "make it up to them" when they argued cases in his courtroom. While he did not guarantee that they would win those cases, he did promise to give extra considerations to their arguments if they made significant donations to his campaign.

End of Block 5: II - Experimental Vignette Part 1/3

Page Break

Start of Block 6: II - Experimental Vignette Part 2/3

[Each subject was randomly assigned to one of the following texts for part 2/3 of the total vignette. All conditions in this part were assigned with equal probability. Assignments of conditions in this part were independent of assignments in other vignette parts.]

[K. High Average Donation]

The average donation amount resulting from these efforts was $1,010 per donor.

[L. Medium Average Donation]

The average donation amount resulting from these efforts was $480 per donor.

[M. Low Average Donation]

The average donation amount resulting from these efforts was $45 per donor.

[N. High Total Donations]

The total donation amount resulting from these efforts was $101,000.

[O. Medium Total Donations]

The total donation amount resulting from these efforts was $48,000.

[P. Low Total Donations]

The total donation amount resulting from these efforts was $4,500.

End of Block 6: II - Experimental Vignette Part 2/3

Page Break

Start of Block 7: II - Experimental Vignette Part 3/3

[Each subject was randomly assigned to one of the following texts for part 3/3 of the total vignette. All conditions in this part were assigned with equal probability. Assignments of conditions in this part were independent of assignments in other vignette parts.]

[Q. Recusal - No Thank You Letter]

Anderson ended up winning the election and is now serving as a judge for the Fairview County Trial Court.

When he later realized that some of those that donated to his campaign would end up as parties or attorneys in his court, he decided to remove himself from any cases involving a donor (the cases were reassigned to other judges).

[R. No Recusal - No Thank You Letter]

Anderson ended up winning the election and is now serving as a judge for the Fairview County Trial Court.

Although he later realized that some of those that donated to his campaign would end up as parties or attorneys in his court, he decided not to remove himself from any cases involving a donor (the cases were not reassigned to other judges).

[S. Recusal Not Considered - No Thank You Letter]

Anderson ended up winning the election and is now serving as a judge for the Fairview County Trial Court.

[T. Recusal - Thank You Letter]

Anderson ended up winning the election and is now serving as a judge for the Fairview County Trial Court.

To express his gratitude for those who helped him win his election, Anderson sent personally-

signed thank you letters to every individual who donated to his campaign.

When he later realized that some of those that donated to his campaign would end up as parties or attorneys in his court, he decided to remove himself from any cases involving a donor (the cases were reassigned to other judges).

[U. No Recusal - Thank You Letter]

Anderson ended up winning the election and is now serving as a judge for the Fairview County Trial Court.

To express his gratitude for those who helped him win his election, Anderson sent personally-signed thank you letters to every individual who donated to his campaign.

Although he later realized that some of those that donated to his campaign would end up as parties or attorneys in his court, he decided not to remove himself from any cases involving a donor (the cases were not reassigned to other judges).

[V. Recusal Not Considered - Thank You Letter]

Anderson ended up winning the election and is now serving as a judge for the Fairview County Trial Court.

To express his gratitude for those who helped him win his election, Anderson sent personally-signed thank you letters to every individual who donated to his campaign.

End of Block 7: II - Experimental Vignette Part 2/3

Page Break

Start of Block 8: III - Introduction

Now we would like to ask you some questions about the description you just read. Please tell us which answer comes closest to how you feel.

End of Block 8: III - Introduction

Page Break

Start of Block 9: III - Donation Questions

[The questions presented in this section vary slightly depending on which text was previously assigned to the subject in part 1/3 of the vignette section.]

[If "A" was assigned]

Imagine that you lived in Fairview County during the judicial election. You hear about Anderson's election campaign. How likely would you be to donate to his campaign?
- Extremely likely
- Moderately likely
- Slightly likely
- Neither likely nor unlikely
- Slightly unlikely
- Moderately unlikely
- Extremely unlikely

[If "B" was assigned]

Imagine that you were an attorney in Fairview County during the judicial election. You receive the letter from Anderson asking for donations. How likely would you be to donate to his campaign?
- Extremely likely
- Moderately likely
- Slightly likely
- Neither likely nor unlikely
- Slightly unlikely
- Moderately unlikely
- Extremely unlikely

[If "C" was assigned]

Imagine that you were a citizen of Fairview County during the judicial election. You receive the letter from Anderson asking for donations. How likely would you be to donate to his campaign?
- Extremely likely
- Moderately likely
- Slightly likely
- Neither likely nor unlikely
- Slightly unlikely
- Moderately unlikely
- Extremely unlikely

[If "D" was assigned]

Imagine that you were an attorney in Fairview County during the judicial election. You receive the letter from Anderson's campaign committee asking for donations. How likely would you be to donate to his campaign?
- Extremely likely
- Moderately likely
- Slightly likely
- Neither likely nor unlikely
- Slightly unlikely
- Moderately unlikely

- Extremely unlikely

[If "E" was assigned]

Imagine that you were a citizen of Fairview County during the judicial election. You receive the letter from Anderson's campaign committee asking for donations. How likely would you be to donate to his campaign?
- Extremely likely
- Moderately likely
- Slightly likely
- Neither likely nor unlikely
- Slightly unlikely
- Moderately unlikely
- Extremely unlikely

[If "F" was assigned]

Imagine that you were an attorney in Fairview County during the judicial election. You attend the donation event, and Anderson asks you to donate. How likely would you be to donate to his campaign?
- Extremely likely
- Moderately likely
- Slightly likely
- Neither likely nor unlikely
- Slightly unlikely
- Moderately unlikely
- Extremely unlikely

[If "G" was assigned]

Imagine that you were a citizen of Fairview County during the judicial election. You attend the donation event, and Anderson asks you to donate. How likely would you be to donate to his campaign?
- Extremely likely
- Moderately likely
- Slightly likely
- Neither likely nor unlikely
- Slightly unlikely
- Moderately unlikely
- Extremely unlikely

[If "H" was assigned]

Imagine that you were an attorney in Fairview County during the judicial election. You attend the donation event, and Anderson's campaign committee asks you to donate. How likely would you be to donate to Anderson's campaign?
- Extremely likely

- Moderately likely
- Slightly likely
- Neither likely nor unlikely
- Slightly unlikely
- Moderately unlikely
- Extremely unlikely

[If "I" was assigned]

Imagine that you were a citizen of Fairview County during the judicial election. You attend the donation event, and Anderson's campaign committee asks you to donate. How likely would you be to donate to his campaign?
- Extremely likely
- Moderately likely
- Slightly likely
- Neither likely nor unlikely
- Slightly unlikely
- Moderately unlikely
- Extremely unlikely

[If "J" was assigned]

Imagine that you are an attorney in Fairview County during the election. You are one of the attorneys from whom Anderson asked for donations. How likely would you be to donate to his campaign?
- Extremely likely
- Moderately likely
- Slightly likely
- Neither likely nor unlikely
- Slightly unlikely
- Moderately unlikely
- Extremely unlikely

End of Block 9: III - Donation Questions

Page Break

Start of Block 10: III - Impartiality Question

Do you believe Judge Anderson can serve as a fair and impartial judge for Fairview County?
- I strongly believe Anderson can be fair and impartial
- I somewhat believe Anderson can be fair and impartial
- I somewhat believe Anderson cannot be fair and impartial
- I strongly believe Anderson cannot be fair and impartial

End of Block 10: III - Impartiality Question

Page Break

Start of Block 11: III - Legitimacy Question

Assume for the moment that all elected judges in the Fairview County Court used the same campaigning strategies as Judge Anderson. How legitimate would you consider the Fairview County Court?

- I would consider it a very legitimate institution
- I would consider it a somewhat legitimate institution
- I would not consider it a very legitimate institution
- I would not consider it a very legitimate institution at all

End of Block 11: III - Legitimacy Question

Page Break

Start of Block 12: III - Accept Decisions As Citizen Question

How likely would you be to accept decisions made by Judge Anderson as impartial, fair, and legitimate?

- Extremely likely
- Moderately likely
- Slightly likely
- Neither likely nor unlikely
- Slightly unlikely
- Moderately unlikely
- Extremely unlikely

End of Block 12: III - Accept Decisions As Citizen Question

Page Break

Start of Block 13: III - Judicial Bias Questions

Assume for a moment that one of the individuals who donated to Judge Anderson is now an attorney who is arguing a case in Judge Anderson's Court. Do you think that the relationship between the donor and Judge Anderson will bias Judge Anderson's judicial behavior?

- It will definitely bias the judge's behavior.
- It will probably bias the judge's behavior.
- It will probably not bias the judge's behavior.

- It will definitely not bias the judge's behavior.

Page Break

Assume for a moment that one of the individuals who donated to Judge Anderson is now a party in a case in Judge Anderson's Court. Do you think that the relationship between the donor and Judge Anderson will bias Judge Anderson's judicial behavior?

- It will definitely bias the judge's behavior.
- It will probably bias the judge's behavior.
- It will probably not bias the judge's behavior.
- It will definitely not bias the judge's behavior.

End of Block 13: III - Judicial Bias Questions

Page Break

Start of Block 14: III - State Election Law Knowledge Quiz

Election Law Which of the following laws reflects what you understand the actual election laws in your state to be? Please select all that apply.

- Judges running for election are allowed to ask for campaign donations.
- Judges running for election are not allowed to ask for campaign donations, but their campaign committees can ask for donations.
- Judges running for election are not allowed to receive any donations.
- If a judge is assigned to case in which a party or an attorney has donated to the judge's political campaign, the judge has to remove himself or herself from the case, and the case will be assigned to another judge.
- Judges do not run for election in my state.

End of Block 14: III - State Election Law Knowledge Quiz

Page Break

Start of Block 15: IV - Introduction

Finally, we have a question about the hypothetical description of Judge Anderson you read earlier. Please answer to the best of your ability.

End of Block 15: IV - Introduction

Page Break

Start of Block 16: IV - Manipulation Check

[The questions presented in this section vary slightly depending on which text was previously assigned to the subject in part 1/3 of the vignette section. The correct answer will also depend on that previous assignment.]

[If "B", "C", "D", or "E" was assigned]

In the vignette you read, who sent the campaign letters?
- Judge Anderson sent out his own letters asking for campaign donations.
- Judge Anderson's Campaign Committee sent out letters asking for campaign donations on Judge Anderson's behalf.
- The Florida State Election Bureau sent out letters asking for campaign donations on Judge Anderson's behalf.
- No one sent letters asking for campaign donations.

[If "F", "G", "H", or "I" was assigned]

In the vignette you read, who asked potential donors for campaign donations at the fundraising event?
- Judge Anderson asked for campaign donations.
- Judge Anderson's Campaign Committee asked for campaign donations on Judge Anderson's behalf.
- The Florida State Election Bureau asked for campaign donations on Judge Anderson's behalf.
- No one asked for campaign donations.

[If "A" or "J" was assigned]

In the vignette you read, who asked potential donors for money?
- Judge Anderson
- Judge Anderson's Campaign Committee
- The Florida State Election Bureau
- No one asked for donations

End of Block 16: V - Manipulation Check

Page Break

Start of Block 17: Ending

Thank you for participating in this survey. Did you run into any problems or errors in the survey? Was there anything confusing? Please explain below.

_____

End of Block 17: Ending

END OF SURVEY